METHODS OF LOGIC

160
Q4m

METHODS OF LOGIC

FOURTH EDITION

W. V. QUINE

Harvard University Press
Cambridge, Massachusetts
1982

74300

Copyright 1950, © 1959, 1972, 1978, 1982 by W. V. Quine
All rights reserved
Printed in the United States of America

Library of Congress Cataloging in Publication Data

Quine, W. V. (Willard Van Orman)
Methods of logic, fourth edition.

Bibliography: p.
Includes index.
1. Logic. I. Title.
BC71.Q5 1982 160 81–22929
ISBN 0–674–57175–4 AACR2
ISBN 0–674–57176–2 (pbk.)

TO MARJORIE

PREFACE

This book undertakes both to convey a precise understanding of the formal concepts of modern logic and to develop convenient techniques of formal reasoning. In each successive edition the book has grown. The portions that can be passed over without detriment to an understanding of basic logical theory amount now to about a third of the book. I have bracketed them so that they can conveniently be skipped, entirely or selectively, for purposes of a shorter course. But in my own teaching I have always covered the whole book, in its past editions, in a semester course.

In the second edition, 1959, the main improvements were two: a more readable proof of the soundness of my system of natural deduction and, in an appendix, proofs of completeness and Löwenheim's theorem.

The third edition, 1972, was a more drastic departure: a new book by half. Proof procedures and decision procedures proliferated, offering a sheaf of alternatives. Approaching logical structure thus from a plurality of angles conduced, I felt, to depth of understanding.

My main method, as I called it, in quantification theory was remarkably simple and easily justified and it was strategic as a base from which to develop the array of alternative methods. Also it admitted of a relatively easy completeness proof. Accordingly my system of natural deduction, which had dominated the first two editions, was demoted in the third edition and accorded only brief treatment in the omissible sector. Responding now to teachers who liked that system and regretted its demotion, I have in this new edition restored it to full treatment, though still making it omissible.

Another major departure in that third edition was the postponement of quantifiers by developing monadic logic first along Boolean lines, though without assuming classes. In this new edition I have pressed the idea further, by exploiting also the set-theoretic notation of class abstraction without assuming classes. Abstracts emerge with all the innocence of relative clauses. The device pays its way in the treatment of substitution, it clarifies the pronominal character of bound variables, and it eventuates in a philosophically agreeable perspective on classes. Further illumination

of the bound variable is undertaken in an omissible chapter on combinatory functors.

There have been patches where points explained in my earlier books needed to be explained again in essentially the same old way. At these points I have adapted examples and expository passages from *Mathematical Logic, Elementary Logic,* and *O Sentido da nova lógica,* preferring not to obscure genuine points of contact by *ad hoc* shifts of examples or of phrasing. But the places are few. Chapter 1 draws in part on §6 of *O Sentido* and Chapter 5 draws in part on §2 of *Mathematical Logic.* In Chapters 4, 8, 14, and 31, examples are borrowed from *Elementary Logic* but are handled differently.

Out of loyalty to loyal teachers of this book, I have been reluctant to depart from the notations of truth functions and quantifiers that I inherited from Whitehead and Russell. In this edition at last I do so. I switch to other symbols that are increasingly in vogue for the conditional, the biconditional, and universal quantification. I make the move not just as a response to fashion, but for technical reasons that will be noted as the symbols emerge.

Partly because of the changed notation, the copy-editing of this edition was unusually exacting. I am grateful to Katarina Rice of Harvard University Press for a perceptive and meticulous job.

CONTENTS

Introduction *1*

PART I. TRUTH FUNCTIONS

1. Negation, Conjunction, and Alternation *9*
2. Truth Functions *16*
3. The Conditional *21*
4. Grouping *27*
5. Truth-Value Analysis *33*
6. Consistency and Validity *40*
7. Implication *45*
8. Words into Symbols *53*
9. Equivalence *60*
10. Alternational Normal Schemata *68*
11. Simplification *74*
12. Duality *79*
13. Axioms *85*

PART II. GENERAL TERMS AND QUANTIFIERS

14. Categorical Statements *93*
15. Venn's Diagrams *98*
16. Syllogisms *102*
17. Limits of These Methods *109*
18. Boolean Schemata *114*
19. Tests of Validity *121*
20. Some Boolean Incidentals *128*
21. The Bound Variable *132*
22. Quantification *136*
23. Rules of Passage. Monadic Schemata *142*
24. Prenexity and Purity *148*
25. Validity Again *154*
26. Substitution *160*

PART III. GENERAL THEORY OF QUANTIFICATION

27. Schemata Extended *167*
28. Substitution Extended *174*
29. Pure Existentials *182*
30. The Main Method *190*
31. Application *195*
32. Completeness *203*
33. Löwenheim's Theorem *209*
34. Decisions and the Undecidable *213*
35. Functional Normal Forms *219*
36. Herbrand's Method *224*
37. Other Methods for Validity *232*
38. Deduction *237*
39. Soundness *245*
40. Deductive Strategy *248*

PART IV. GLIMPSES BEYOND

41. Singular Terms *259*
42. Identity *268*
43. Descriptions *274*
44. Elimination of Singular Terms *278*
45. Elimination of Variables *283*
46. Classes *288*
47. Number *294*
48. Axiomatic Set Theory *300*

Partial Answers to Exercises *305*
Bibliography *320*
Index *327*

INTRODUCTION

Logic, like any science, has as its business the pursuit of truth. What are true are certain statements; and the pursuit of truth is the endeavor to sort out the true statements from the others, which are false.

Truths are as plentiful as falsehoods, since each falsehood admits of a negation which is true. But scientific activity is not the indiscriminate amassing of truths; science is selective and seeks the truths that count for most, either in point of intrinsic interest or as instruments for coping with the world.

Strictly speaking, what admit of truth and falsity are not statements as repeatable patterns of utterance, but individual events of statement utterance. For, utterances that sound alike can vary in meaning with the occasion of the utterance. This is due not only to careless ambiguities, but to systematic ambiguities which are essential to the nature of language. The pronoun 'I' changes its reference with every change of speaker; 'here' changes its reference with every significant movement through space; and 'now' changes its reference every time it is uttered.

The crucial point of contact between description and reality is to be sought in the utterance of a statement on the occasion of a stimulation to which that string of words has become associated. Not that the statement will refer to stimulation or sensation; it is apt to refer to physical objects. Language is a social institution, serving, within its limitations, the social end of communication; so it is not to be wondered that the objects of our first and commonest utterances are socially shared physical objects rather than private experiences. Physical objects, if they did not exist, would (to transplant Voltaire's epigram) have had to be invented. They are indispensable as the public common denominators of private sense experience.

The latest scientific pronouncement about positrons and the statement that my pen is in my hand are equally statements about physical objects; and physical objects are known to us only as parts of a systematic conceptual structure which, taken as a whole, impinges at its edges upon observation. We have a network of statements that are variously linked with one another and some of which, out at the periphery of the network, are associated more or less strongly with sensory stimulation. Even the peripheral ones are mostly about physical objects: examples are 'My pen is in my hand', 'The mercury is at 80'.

A sensory stimulation elicits some closely associated statement and the associations then reverberate through the system of statements, activating at length another peripheral statement whose sensory associations make us expect some particular further stimulation. Such, schematically, is the mechanism of prediction. When prediction fails, we question the intervening network of statements. We retain a wide choice as to what statements of the system to preserve and what ones to revise; any one of many revisions will suffice to unmake the particular implication that brought the system to grief.

Normally the peripheral statements, closely associated to stimulations, are to be preserved from revision once the appropriate stimulations have occurred. If revision of the system should become necessary, other statements than these are to suffer. It is only by such an allocation of priority that we can hope to claim any empirical content or objective reference for the system as a whole.

There is also, however, another and somewhat opposite priority: the more fundamental a law is to our conceptual scheme, the less likely we are to choose it for revision. When some revision of our system of statements is called for, we prefer, other things being equal, a revision which disturbs the system least.

Where the two priorities come into conflict, either is capable of prevailing. Statements close to experience and seemingly verified by the appropriate experiences may occasionally be given up, even by pleading hallucination, in the extreme case where their retention would entail a cataclysmic revision of fundamental laws. But to overrule a multiplicity of such statements, if they reinforce one another and are sustained by different observers, would invite criticism.

The priority on law, considered now apart from any competition with the priority on statements verified by experience, admits of many gradations. Conjectures of history and economics will be revised more willingly than laws of physics, and these more willingly than laws of mathematics and logic. Our system of statements has such a thick cushion of indeterminacy, in relation to experience, that vast domains of law can easily be held immune to revision on principle. We can always turn to other quarters of the system when revisions are called for by unexpected experiences. Mathematics and logic, central as they are to the conceptual scheme, tend to be accorded such immunity, in view of our conservative preference for revisions which disturb the system least; and herein, perhaps, lies the "necessity" which the laws of mathematics and logic are felt to enjoy.

In the end it is perhaps the same to say, as one often does, that the laws of mathematics and logic are true simply by virtue of our conceptual scheme. For, it is certainly by virtue of that scheme that those laws are central to it; and it is by virtue of being thus central that the laws are preserved from revision at the expense of statements less strategically situated.

But it must now be remarked that our conservative preference for those revisions which disturb the system least is opposed by a significant contrary force, a force for simplification. Far-reaching revision of the fundamental laws of physics was elected in recent decades, by considerations of simplicity, in preference to the welter of *ad hoc* subsidiary laws which would otherwise have been needed to accommodate the wayward experiences of Michelson and Morley and other experimenters. Continued experiment "confirmed" the fundamental revisions, in the sense of increasing the simplicity differential.

Mathematical and logical laws themselves are not immune to revision if it is found that essential simplifications of our whole conceptual scheme will ensue. There have been suggestions, stimulated largely by quandaries of modern physics, that we revise the true-false dichotomy of current logic in favor of some sort of tri- or n-chotomy. Logical laws are the most central and crucial statements of our conceptual scheme, and for this reason the most protected from revision by the force of conservatism; but, because again of their crucial position, they are the laws an apt revision of which might offer the most sweeping simplification of our whole system of knowledge.

Thus the laws of mathematics and logic may, despite all "necessity," be abrogated. But this is not to deny that such laws are true by virtue of the conceptual scheme, or by virtue of meanings. Because these laws are so central, any revision of them is felt to be the adoption of a new conceptual scheme, the imposition of new meanings on old words. No such revolution, by the way, is envisaged in this book; there will be novelties of approach and technique in these pages, but at bottom logic will remain unchanged.

I have been stressing that in large part our statements are linked only remotely to observation. It is only by way of the relations of one statement to another that the statements in the interior of the system can figure at all in the prediction of experience, and can be found deserving of revision when prediction fails. Now of these relations of statements to statements, one of conspicuous importance is the relation of logical implication: the relation of any statement to any that follows logically from

it. If one statement is to be held as true, each statement implied by it must also be held as true; and thus it is that statements internal to the system have their effects on statements at the periphery.

But for implication, our system of statements would for the most part be meaningless; nothing but the periphery would make sense. Yet implication is not really an added factor; for, to say that one statement logically implies a second is the same as saying that a third statement of the system, an 'if-then' compound formed from the other two, is logically true or "valid." Logical truths are statements on a par with the rest, but very centrally situated; they are statements of such forms as '*p* or not *p*', 'If *p* then *p*', 'If *p* and *q* then *q*', 'If everything is thus and so then something is thus and so', and others more complex and less quickly recognizable. Their characteristic is that they not only are true but stay true even when we make substitutions upon their component words and phrases as we please, provided merely that the so-called "logical" words '=', 'or', 'not', 'if-then', 'everything', 'something', etc., stay undisturbed. We may write any statements in the '*p*' and '*q*' positions and any terms in the 'thus and so' positions, in the forms cited above, without fear of falsity. All that counts, when a statement is logically true, is its structure in terms of logical words. Thus it is that logical truths are commonly said to be true by virtue merely of the meanings of the logical words.

The chief importance of logic lies in implication, which, therefore, will be the main theme of this book. Techniques are wanted for showing, given two statements, that the one implies the other; herein lies logical deduction. Such techniques will be developed, for increasingly inclusive portions of logic, as the book proceeds. The objects of deduction, the things related by implication, are statements; so statements will constitute not merely the medium of this book (as of most), but the primary subject matter.

Strictly speaking, as urged earlier, what admit of truth and falsity are not the statements but the individual events of their utterance. However, it is a great source of simplification in logical theory to talk of statements in abstraction from the individual occasions of their utterance; and this abstraction, if made in full awareness and subject to a certain precaution, offers no difficulty. The precaution is merely that we must not apply our logical techniques to examples in which one and the same statement recurs several times with changed meanings, owing to variations in immediate context. But such examples are easily enough adjusted to the purposes of logic by some preliminary paraphrasing, by way of bringing the implicit shifts of meaning into explicit form. (See Chapter 8.)

Logic and mathematics were coupled, in earlier remarks, as jointly enjoying a central position within the total system of discourse. Logic as commonly presented, and in particular as it will be presented in this book, seems to differ from mathematics in that in logic we talk about statements and their interrelationships, notably implication, whereas in mathematics we talk about abstract nonlinguistic things: numbers, functions, and the like. This contrast is in large part misleading. Logical truths, e.g., statements of the form 'If p and q then q', are not about statements; they may be about anything, depending on what statements we put in the blanks 'p' and 'q'. When we talk *about* such logical truths, and when we expound implications, we are indeed talking about statements; but so are we when we talk *about* mathematical truths.

But it is indeed the case that the truths of mathematics treat explicitly of abstract nonlinguistic things, e.g., numbers and functions, whereas the truths of logic, in a reasonably limited sense of the word 'logic', have no such entities as specific subject matter. This is an important difference. Despite this difference, however, logic in its higher reaches is found to bring us by natural stages into mathematics. For, it happens that certain unobtrusive extensions of logical theory carry us into a realm, sometimes also called 'logic' in a broad sense of the word, which does have abstract entities of a special kind as subject matter. These entities are classes; and the logical theory of classes, or set theory, proves to be the basic discipline of pure mathematics. From it, as first came to be known through the work of Frege, Dedekind, Weierstrass, and their successors in the late nineteenth century and after, the whole of classical mathematics can be generated. Before the end of the book we shall have ascended through four grades of logic in the narrower sense, and emerged into set theory; and here we shall see, as examples of the derivation of classical mathematics, how the concept of number and various related notions can be defined.

I
TRUTH FUNCTIONS

1

NEGATION, CONJUNCTION, AND ALTERNATION

The peculiarity of *statements* which sets them apart from other linguistic forms is that they admit of truth and falsity, and may hence be significantly affirmed and denied. To deny a statement is to affirm another statement, known as the *negation* or *contradictory* of the first. To deny 'The Taj Mahal is white' is to affirm 'The Taj Mahal is not white'. Note that this negation is opposed to the original not as black to white, but as nonwhite to white; it counts as true in every case except the case of whiteness.

The commonest method of forming the negation of statements in ordinary language is by attaching 'not' (or 'does not', etc.) to the main verb, as in the foregoing example. But if the verb is governed by 'sometimes' or 'always', the negation is formed rather by substituting 'never', or 'not always'. If the statement is compound and thus has no main verb, its negation has to be phrased more elaborately; e.g., 'It is not the case both that. . . and that. . .'. But, despite such irregularities of ordinary language, a little care suffices for constructing a clear negation of any given statement, the guiding consideration being simply this: the negation is to count as false if the given statement is true, and the negation is to count as true under any and all circumstances under which the given statement is false.

In logical studies it is convenient to adopt a single sign of negation, consisting of the prefix '$-$', applied to statements as wholes. Thus '$-$(Jones is away)' means 'Jones is not away'; the parentheses here serve to group, as a single whole, the statement to which '$-$' is applied. The sign '$-$' might be translated into words as 'it is not the case that'; briefly it may be pronounced 'not'. When a statement is represented as a single letter 'p', as is commonly done in logical discussion, the sign of negation will be placed above instead of in front; thus we shall write '$\bar{p}$' instead of '$-p$' for the negation of 'p'.

Instead of affirming each of several statements we can, equivalently, affirm a single statement which is known to logicians (in contrast to grammarians) as the *conjunction* of the given statements. The conjunction of two or more statements is commonly expressed in English by linking

the statements by 'and', or commas, or a combination of the two: 'Some are born great, some achieve greatness, and some have greatness thrust upon them.' In logical studies it is convenient to express the conjunction simply by writing the component statements in juxtaposition; e.g., '(some are born great) (some achieve greatness) (some have greatness thrust upon them)'—where again the parentheses serve mainly to mark off the component statements as wholes. If we think of 'p', 'q', and 'r' as statements, their conjunction is represented as 'pqr'.

The meanings of negation and conjunction are summed up in these laws: *The negation of a true statement is false; the negation of a false statement is true; a conjunction of statements all of which are true is true; and a conjunction of statements not all of which are true is false.*

We see immediately that '$\bar{\bar{p}}$', the negation of '$\bar{p}$', will be true if and only if '$\bar{p}$' is false, hence if and only if 'p' is true; so there is no point in writing a double negation '$\bar{\bar{p}}$', amounting as it does simply to 'p'. It is equally evident that the conjunction 'pp' amounts simply to 'p'.

Consider now '$p(qr)$'. This, being the conjunction of 'p' and 'qr', is to be true if and only if 'p' and 'qr' are both true; and 'qr' in turn is to be true if and only if 'q' and 'r' are both true. Hence '$p(qr)$' is true if and only if 'p', 'q', and 'r' are all true; in other words, '$p(qr)$' amounts simply to the three-way conjunction 'pqr'. In the same way it may be seen that '$(pq)r$' amounts simply to 'pqr'. We may therefore drop parentheses and always write 'pqr', viewing this at will as the conjunction of 'pq' and 'r', as the conjunction of 'p' and 'qr', and as the conjunction of 'p', 'q', and 'r'. Conjunction is, in mathematical jargon, *associative:* internal grouping is immaterial in 'pqr', just as in the sum '$x + y + z$' or product 'xyz' of arithmetic. Conjunction contrasts in this respect with the arithmetical operation of division; for note that the parentheses in '$12 \div (6 \div 2)$' and '$(12 \div 6) \div 2$' make all the difference between 4 and 1.

While the associative law for addition, multiplication, and conjunction justifies suppressing parentheses, their suppression should not make us forget that the law is still there and doing its work. We may best regard the notation 'pqr' of three-way conjunction as an abbreviation specifically and arbitrarily of (say) '$(pq)r$'; then, when in practice we treat 'pqr' also as '$p(qr)$', what is strictly afoot rather is a logical transformation of '$(pq)r$' into its equivalent '$p(qr)$'.

Another respect in which conjunction resembles addition and multiplication, and differs from division, is that it is *commutative;* i.e., order is immaterial, there being no need to distinguish between 'pq' and 'qp'.

But conjunction was lately observed to enjoy also a third convenient property, not shared by addition and multiplication; viz., '*pp*' reduces to '*p*'. Conjunction is *idempotent,* to persist in the jargon. Taken together, these three properties of conjunction come simply to this: once we have an inventory of all the distinct components of a continued conjunction, no further details of the constitution of the conjunction need concern us.

Having touched on negation and conjunction, which correspond to 'not' and 'and', we turn now to a third way of forming statements from statements. It is called *alternation,* and corresponds to the connective 'or', or 'either-or'. This connective is subject in ordinary discourse to conflicting usages. One sense is the *nonexclusive,*[1] according to which the compound is true so long as at least one of the components is true. Under this usage the statement:

(Either) Jones is ill or Smith is away

is true if Jones is ill and Smith is away, true again if Jones is not ill but Smith is away, true again if Jones is ill but Smith is not away, and false only in case Jones is neither ill nor Smith away. The other sense in which 'or' is sometimes used, called the *exclusive,* construes the compound as true just in case *exactly* one of the components is true. In this sense of 'or', the compound becomes false not only when the components are both false (Jones neither ill nor Smith away) but also when the components are both true (Jones ill and Smith away).

The ambiguity of 'or' is commonly resolved, in ordinary usage, by adding the words 'or both' or 'but not both'. Thus the nonexclusive sense is expressible in the unambiguous fashion:

Jones is ill or Smith is away or both,

and the exclusive sense thus:

Jones is ill or Smith is away but not both.

When we are confronted with '*p* or *q*' by itself, we do not in general know which interpretation to assign to it. Often the choice is immaterial, in that either sense would serve equally. For example, consider the ex-

[1] I follow Cooley in preferring this awkward term to the more usual but somewhat misleading 'inclusive'.

pression '$x \leqq y$', i.e., '$x < y$ or $x = y$'. It makes no difference whether 'or' here is understood in the nonexclusive or the exclusive sense. The only difference between the two senses occurs in the case where both components are true; but when the components concerned are '$x < y$' and '$x = y$', the case of joint truth does not arise either in fact or in the mind of the speaker.

It is a common error to believe that examples like '$x < y$ or $x = y$' are clear cases of the use of 'or' in the exclusive sense, and in consequence of this error there is a tendency to overestimate the role which the exclusive sense of 'or' plays in everyday language. The clauses '$x < y$' and '$x = y$' are, of themselves, mutually exclusive or incompatible clauses; but this incompatibility, far from establishing that the context '$x < y$ or $x = y$' uses 'or' in the exclusive sense, deprives us of the one case in which we might hope to distinguish between the exclusive and nonexclusive senses. Since the clauses '$x < y$' and '$x = y$' are already of such nature as to exclude each other, it is immaterial whether we understand 'or' as repeating this exclusion or not.

If we want to establish indisputable instances of the exclusive use of 'or', we must imagine circumstances in which the person who uses 'or' has a positive purpose of denying, explicitly within the given statement, the joint truth of the components. Such examples are rare, but they exist. In an example given by Tarski it is supposed that a child asks his father to take him to the beach and afterwards to the movie. The father replies, in a tone of refusal, "We will go either to the beach or to the movie." Here the exclusive use is clear; the father means simultaneously to promise and to refuse. But it is much easier to find cases in which the nonexclusive interpretation is obligatory. For example, when it is decreed that passports will be issued only to persons who were born in the country or who are married to natives of the country, this does not mean that passports will be refused to persons who were born in the country and are married to natives. Most use of 'or' in everyday language is either of this type which admits only of the nonexclusive interpretation, or of the type of '$x < y$ or $x = y$', which admits both interpretations indifferently.

Latin has distinct words for the two senses of 'or': *vel* for the nonexclusive and *aut* for the exclusive. In modern logic it is customary to write '$\vee$', reminiscent of 'vel', for 'or' in the nonexclusive sense: '$p \vee q$'. It is this mode of compounding statements, and only this, that is called *alternation*. When the ambiguous 'or' of ordinary language appears hereafter in the book, let us agree to construe it in this nonexclusive sense. If

occasions arise where the exclusive sense of 'or' is really wanted, it is easy enough to express it explicitly:

p or q but not both,

or equivalently:

Either p and not q or q and not p,

i.e., in symbols:

$$p\bar{q} \vee \bar{p}q.^{2}$$

The meaning of alternation, then, is given by this rule: *An alternation is true if at least one of the components is true, and otherwise false.* Whereas a conjunction is true if and only if its components are all true, an alternation is false if and only if its components are all false. In a metaphor from genetics, conjunction and alternation may be contrasted thus: in conjunction, truth is recessive and falsity dominant; in alternation, truth is dominant and falsity recessive.

Because the explanation of alternation is just the same as that of conjunction except for interchanging the roles of truth and falsehood, it is evident that the formal properties of conjunction must reappear as properties of alternation; thus alternation, like conjunction, is associative, commutative, and idempotent. We can render '$(p \vee q) \vee r$' and '$p \vee (q \vee r)$' indifferently as '$p \vee q \vee r$'; we can interchange '$p \vee q$' with '$q \vee p$'; and we can reduce '$p \vee p$' to 'p'. All that matters in a continued alternation, as in a continued conjunction, is an inventory of the distinct components.

Though the grouping of components is irrelevant within a continued conjunction and within a continued alternation, it is important where conjunction and alternation are mixed; we must distinguish, e.g., between '$pq \vee r$' and '$p(q \vee r)$'. In Chapter 5 a systematic technique will appear whereby all complexes of conjunction, alternation, and negation can conveniently be analyzed; meanwhile, however, it is easy to see in advance that '$pq \vee r$' and '$p(q \vee r)$' are bound to behave in quite unlike ways. One clear point of divergence is this: '$p(q \vee r)$', being a conjunction with 'p' as a component, cannot be true unless 'p' is true, whereas

[2] After Chapter 3, this may be also written '$p \leftrightarrow \bar{q}$'.

'$pq \vee r$', being an alternation with 'r' as one component, will be true so long as 'r' is true, even if 'p' be false.

Grouping is likewise important when negation occurs in combination with conjunction or alternation. We are not likely, indeed, to confuse '$\bar{p}q$' with '$-(pq)$', nor '$\bar{p} \vee q$' with '$-(p \vee q)$', for in the one case only 'p' is negated while in the other case the whole compound is negated. But what is less evident is that we must distinguish also between '$-(pq)$' and '$\bar{p}\bar{q}$', and between '$-(p \vee q)$' and '$\bar{p} \vee \bar{q}$'. Let us see what these distinctions are, taking 'p' as 'penicillin was flown in' and 'q' as 'a quarantine was imposed'. There are four possible situations:

pq: Penicillin was flown in and a quarantine was imposed.
$\bar{p}q$: Penicillin was not flown in but a quarantine was imposed.
$p\bar{q}$: Penicillin was flown in and no quarantine was imposed.
$\bar{p}\bar{q}$: Penicillin was not flown in nor was a quarantine imposed.

Now '$-(pq)$' denies just the first of the four situations, and so comes out true in the second, third, and fourth. Thus '$-(pq)$' is quite different from '$\bar{p}\bar{q}$', which holds in the fourth case only. As for '$\bar{p} \vee \bar{q}$', this holds whenever one or both of '$\bar{p}$' and '$\bar{q}$' hold; hence in the second, third, and fourth cases. We can therefore equate '$\bar{p} \vee \bar{q}$' with '$-(pq)$'. Finally '$-(p \vee q)$' holds in the one case where '$p \vee q$' fails—hence in the fourth case alone; so we may equate '$-(p \vee q)$' with '$\bar{p}\bar{q}$'.

So '$-(pq)$' does not amount to '$\bar{p}\bar{q}$', but to '$\bar{p} \vee \bar{q}$'; and '$-(p \vee q)$' does not amount to '$\bar{p} \vee \bar{q}$', but to '$\bar{p}\bar{q}$'. We may distribute the negation sign of '$-(pq)$' and '$-(p \vee q)$' over 'p' and 'q' individually only on pain of changing conjunction to alternation and vice versa.[3]

A little reflection reveals the same relationship in ordinary language. Clearly '$\bar{p}\bar{q}$', or 'Not p and not q', may be phrased 'Neither p nor q'; and it is scarcely surprising that 'Neither p nor q' should amount to '$-(p \vee q)$', the negation of 'Either p or q'. Again '$-(pq)$' may be read 'Not both p and q', and from this it is no leap to 'Either not p or not q'.

If we read the negation sign as 'it is not the case that', the distinctions of grouping become automatic.

$-(pq)$: It is not the case that both p and q.
$\bar{p}\bar{q}$: It is not the case that p and it is not the case that q.
$-(p \vee q)$: It is not the case that either p or q.
$\bar{p} \vee \bar{q}$: It is not the case that p or it is not the case that q.

[3] These equivalences are called DeMorgan's laws. See Chapter 10.

Of these four the first and last, we have seen, come to the same thing; and similarly for the second and third.

HISTORICAL NOTE: The use of the dash to express negation, and of juxtaposition to express conjunction, comes down from Peano (fl. 1888–1901) and earlier writers. The bar over the letter was used for negation by C. S. Peirce from 1870. The tilde '$\sim$', a modified 'n', saw some use as a negation sign in the nineteenth century and was revived by Whitehead and Russell. It has considerable currency today, also in writings of my own. Because in the present book there is much negating of single letters, however, I have here favored Peirce's bar for its compactness and perspicuity; and, given that, the dash is more in keeping than the tilde as negation sign for longer expressions. The alternation sign '$\vee$' dates from Whitehead and Russell and is now used by almost all logicians. But the term for alternation is less uniform; some writers, following Whitehead and Russell, use the less suggestive term 'disjunction'. Custom varies in the notation for conjunction; some invert the alternation sign for the purpose, and some, following Hilbert (1928), use '&'. For negation there is a deviant sign '$\neg$' that has gained some currency.

EXERCISES

1. Which of the four cases:

   Jones ill, Smith away,
   Jones not ill, Smith away,
   Jones ill, Smith not away,
   Jones not ill, Smith not away

   make the statement:

   Jones is not ill or Smith is not away

   come out true when 'or' is construed exclusively? nonexclusively?

2. Construing 'or' exclusively or nonexclusively is indifferent to the truth or falsity of the statement above in certain of the four cases. Which cases are those?

3. Construing 'p' as 'penicillin was flown in' and 'q' as 'the quarantine was lifted', distinguish in phrasing between '$-(p \vee q)$' and '$\bar{p} \vee q$'.

Under what circumstances would one of these compounds come out true and the other false?

4. Are there circumstances in which '$\bar{p} \vee q$', '$p \vee \bar{q}$', '$\bar{p} \vee \bar{q}$', and '$-(p \vee q)$' all come out true?

2
TRUTH FUNCTIONS

All that is strictly needed for a precise understanding of negation, conjunction, and alternation is stated in these laws:

'$\bar{p}$' is true if and only if 'p' is false,
'$pq \ldots s$' is true if and only if all of 'p', 'q', . . . , 's' are true,
'$p \vee q \vee \ldots \vee s$' is true if and only if 'p', 'q', . . . , 's' are not all false.

Now it is evident from these laws that negation, conjunction, and alternation share the following important property: in order to be able to determine the truth or falsity of a negation, conjunction, or alternation, it is sufficient to know the truth or falsity of the component parts.

It is convenient to speak of truth and falsity as *truth values;* thus the truth value of a statement is said to be truth or falsity according as the statement is true or false. What we have just observed, then, is that the truth value of a negation, conjunction, or alternation is determined by the truth values of its components. This state of affairs is expressed by speaking of negation, conjunction, and alternation as *truth functions*. In general, a compound is called a *truth function of* its components if its truth value is determined in all cases by the truth values of the components. More precisely: a way of forming compound statements from component statements is *truth-functional* if the compounds thus formed always have matching truth values as long as their components have matching truth values.[4]

[4] I thank James A. Thomas for reporting an error in my old phrasing of this.

The property of truth-functionality which is thus enjoyed by negation, conjunction, and alternation may be better appreciated if for contrast we examine a non-truth-functional compound:

Jones died because he ate fish with ice cream.

Even agreeing that the components 'Jones died' and 'Jones ate fish with ice cream' are true, we may still dispute over the truth value of this compound. The truth value of the compound is not determined simply by the truth values of the component statements, but by these in company with further considerations; and very obscure those further considerations are. On the other hand the truth value of the conjunction:

Jones ate fish with ice cream and died

or of the alternation:

Jones ate fish with ice cream or died

or of the negation:

Jones did not die

admits of no dispute whatever once the truth values of 'Jones ate fish with ice cream' and 'Jones died' are known individually.

The compound 'p because q' is shown not to be a truth function of 'p' and 'q' by the fact that it comes out true when some truths are put for 'p' and 'q' and false when other truths are put for 'p' and 'q'. In the case of '$p \vee q$', 'pq', and '$\bar{p}$', on the other hand, one true component is as good as another and one false component is as bad as another so far as the truth or falsity of the compound is concerned.

Any particular truth function can be adequately described by presenting a schedule showing what truth values the compound will take on for each choice of truth values for the components. Our three basic truth functions themselves, indeed, were summarily so described in the opening lines of the present chapter. Any unfamiliar fourth truth-functional symbol could likewise be introduced and adequately explained simply by saying what truth values on the part of the components are to make the new compound true and what ones are to make it false. A symbol

'excl-or' for the exclusive 'or', e.g., would be fully explained by a stipulation that 'p excl-or q' is to be false when 'p' and 'q' are taken as both true or both false, and true in the remaining two cases ('p' true and 'q' false or vice versa).

This question now arises: do our negation, conjunction, and alternation constitute a sufficient language for *all* truth-functional purposes? Given an explanation of a new truth-functional symbol (e.g., 'excl-or'), can we always be sure that the new symbol will be translatable into our existing notation? The answer is that negation and conjunction are always sufficient, without even alternation!

E.g., consider again 'p excl-or q'. This has been explained as false in just the case (a) where 'p' and 'q' are both true and the case (b) where 'p' and 'q' are both false. Therefore 'p excl-or q' amounts simply to denying, simultaneously, 'pq' and '$\bar{p}\bar{q}$'; for 'pq' holds in case (a) and there alone, and '$\bar{p}\bar{q}$' holds in case (b) and there alone. Therefore 'p excl-or q' amounts to:

$$-(pq)\ -(\bar{p}\bar{q}),$$

the conjunction of '$-(pq)$' and '$-(\bar{p}\bar{q})$'; for this conjunction simultaneously denies 'pq' and '$\bar{p}\bar{q}$' and nothing more. The compound 'p excl-or q' is false in the two cases where '$-(pq)\ -(\bar{p}\bar{q})$' is false, and true in the two cases where '$-(pq)\ -(\bar{p}\bar{q})$' is true. So the symbol 'excl-or' is superfluous; conjunction and negation suffice.

In the same way the symbol '$\vee$' of alternation itself can be seen to be superfluous. The one case where '$p \vee q$' is to be false is the case where 'p' and 'q' are both false—i.e., the case where '$\bar{p}\bar{q}$' holds. So instead of writing '$p \vee q$' we may simply deny '$\bar{p}\bar{q}$', writing '$-(\bar{p}\bar{q})$'.

These two simple examples of translating truth functions into negation and conjunction illustrate a general method which works for almost any truth function. Given a description of a truth function—i.e., given simply a schedule showing what truth values the compound is to take on for each choice of truth values for the components—we can construct a truth function out of negation and conjunction which answers the description. The general method will become evident if illustrated once more, this time with a less simple and more arbitrary example than 'excl-or' and '$\vee$'. This time a certain truth function of 'p', 'q', and 'r' is described as follows, let us say. It is to come out true in the five cases:

'p' false, 'q' true, 'r' true,
'p' true, 'q' false, 'r' true,

'p' true, 'q' true, 'r' false,
'p' false, 'q' true, 'r' false,
'p' false, 'q' false, 'r' false

and false in the remaining three cases:

'p' true, 'q' true, 'r' true,
'p' false, 'q' false, 'r' true,
'p' true, 'q' false, 'r' false.

Now these three latter cases are the cases respectively where 'pqr' is true, where '$\bar{p}\bar{q}r$' is true, and where '$p\bar{q}\bar{r}$' is true; so the compound which we are seeking is obtained simply by simultaneously negating these three unwanted cases, in a conjunction thus:

$$-(pqr)\ -(\bar{p}\bar{q}r)\ -(p\bar{q}\bar{r}).$$

Our compound thus denies, explicitly, just those cases in which it was to come out false; in all other cases it comes out true.

Clearly this same method will work for any example so long as there are *some* cases, one or more, in which the desired compound is to come out false. We thus have a routine whereby *almost* any described truth function can be written out in terms of negation and conjunction. The only truth functions which our routine fails to take care of are the ones which are to be true in *all* cases, regardless of the truth values of the components. These trivial exceptions call, then, for separate treatment; and a treatment is straightway forthcoming which is correspondingly trivial. If our problem is to express a truth function of 'p', 'q', 'r', and 's' (say) which will come out true regardless of what truth values are assigned to 'p', 'q', 'r', and 's', we may solve it simply by writing:

$$-(p\bar{p}qrs).$$

Clearly '$p\bar{p}qrs$' will come out false in all cases, on account of '$p\bar{p}$'; therefore '$-(p\bar{p}qrs)$' will come out true in all cases.

So it is now clear that negation and conjunction constitute a sufficient language for all truth-functional purposes. Far from needing ever to add further notations for hitherto inexpressible truth functions, we can even drop the notation '$\vee$' which is already at hand. But we shall not drop it, for it facilitates certain technical manipulations (cf. Chapters 10–12).

It should be remarked that conjunction is really no less superfluous than alternation; for the fact is that an adequate notation for truth functions is constituted not only by negation and conjunction, but equally by negation and alternation. To see this it is sufficient to observe that the conjunction 'pq' itself is translatable into terms of negation and alternation, as '$-(\bar{p} \vee \bar{q})$'. This expression is equivalent to 'pq', in the sense that it comes out true where 'p' and 'q' are both true, and otherwise false. For, '$-(\bar{p} \vee \bar{q})$' is true if and only if '$\bar{p} \vee \bar{q}$' is false, hence if and only if '$\bar{p}$' and '$\bar{q}$' are both false, and hence if and only if 'p' and 'q' are both true.

In lieu of negation and conjunction, or negation and alternation, a *single* connective can be made to suffice—viz., '$|$', construed as follows: '$p \mid q$' is to be true if and only if 'p' and 'q' are not both true. '$p \mid q$' amounts to what would be expressed in terms of conjunction and negation as '$-(pq)$'; but, if we start rather with '$|$' as basic, we can express '$\bar{p}$' in terms of '$|$' as '$p \mid p$', and 'pq' as '$(p \mid q) \mid (p \mid q)$'. Another connective which would suffice by itself is '$\downarrow$', or 'neither-nor'. '$p \downarrow q$' amounts to what would be expressed in terms of conjunction and negation as '$\bar{p}\bar{q}$'; but, if we start rather with '$\downarrow$' as basic, we can express '$\bar{p}$' as '$p \downarrow p$' and 'pq' as '$(p \downarrow p) \downarrow (q \downarrow q)$'.

HISTORICAL NOTE: The logic of alternation, conjunction, and negation was investigated systematically in ancient times by the Stoics,[5] in the Middle Ages by Petrus Hispanus, Duns Scotus, and others,[5] and in modern times mainly by Boole (1847 onward) and Schröder (1877 onward). The terms 'truth value' and 'truth function' are respectively from Frege (1892) and Whitehead and Russell (1910). The reductions to '$|$' and '$\downarrow$' are due to Sheffer (1913).

EXERCISES

1. Obtain a compound of 'p', 'q', and 'r', using only conjunction and negation, which will come out true whenever exactly two of 'p', 'q', and 'r' are true, and otherwise false.

2. Do the same using only alternation and negation. One way would be to translate the previous answer.

[5] See Łukasiewicz, "Zur Geschichte," cited in the Bibliography.

3
THE CONDITIONAL

Besides 'and' and 'or', another connective of statements which plays an important part in everyday language is 'if-then'. A statement of the form 'if p then q' is called a *conditional*. The component in the position of 'p' here is called the *antecedent* of the conditional, and the component in the position of 'q' is called the *consequent*.

A conjunction of two statements is true, we know, just in case both components are true: and an alternation is true just in case one or both components are true. Now under what circumstances is a conditional true? Even to raise this question is to depart from everyday attitudes. An affirmation of the form 'if p then q' is commonly felt less as an affirmation of a conditional than as a conditional affirmation of the consequent.[6] If, after we have made such an affirmation, the antecedent turns out true, then we consider ourselves committed to the consequent, and are ready to acknowledge error if it proves false. If on the other hand the antecedent turns out to have been false, our conditional affirmation is as if it had never been made.

Departing from this usual attitude, however, let us think of conditionals simply as compound statements which, like conjunctions and alternations, admit as wholes of truth and falsity. Under what circumstances, then, should a conditional as a whole be regarded as true, and under what circumstances false? Where the antecedent is true, the above account of common attitudes suggests equating the truth value of the conditional with that of the consequent; thus a conditional with true antecedent and true consequent will count as true, and a conditional with true antecedent and false consequent will count as false. Where the antecedent is false, on the other hand, the adoption of a truth value for the conditional becomes rather more arbitrary; but the decision which proves most convenient is to regard all conditionals with false antecedents as true. The conditional 'if p then q', so construed, is written '$p \rightarrow q$' and called the *material conditional*. It is construed as true where 'p' and 'q' are true, also where 'p' is false and 'q' true, and also where 'p' and 'q' are both false; and it is

[6] I am indebted here to Dr. Philip Rhinelander. Elsewhere in this chapter I draw upon Section 2 of my *Mathematical Logic* by permission of the Harvard University Press.

construed as false only in the remaining case, viz., where 'p' is true and 'q' false.

The sign '$\rightarrow$', like '$\vee$', is superfluous. We know from Chapter 2 how to construct, by means of conjunction and negation alone, a compound which shall be false in just the one case where 'p' is true and 'q' false; viz., '$-(p\bar{q})$'. We could dispense with '$\rightarrow$' altogether, always writing '$-(p\bar{q})$' instead of '$p \rightarrow q$'. Yet another rendering, readily seen to come to the same thing, is '$\bar{p} \vee q$'. However, the superfluous sign '$\rightarrow$' will prove eventually to facilitate technical manipulations.

Now consider the statement:

(1) If anything is a vertebrate, it has a heart.

This, to begin with, is not a conditional in the sense with which we have been concerned above, for it is not really a compound of two statements 'anything is a vertebrate' and 'it has a heart'. The form of words 'it has a heart' is not a statement, true or false, which can be entertained in its own right, and be mooted to be true in case there are vertebrates. Rather, (1) must be viewed as affirming a bundle of individual conditionals: If a is a vertebrate, a has a heart; if b is a vertebrate, b has a heart; and so on. In short:

(2) No matter what x may be, if x is a vertebrate then x has a heart.

But it is important to note that, of the bundle of conditionals which (2) affirms, each individual conditional can quite suitably be interpreted as a material conditional. For, if we reflect that the material conditional '$p \rightarrow q$' amounts to '$-(p\bar{q})$', and then rewrite (2) accordingly, we have:

> No matter what x may be, it is not the case that x both is a vertebrate and does not have a heart.

or briefly:

(3) Nothing is a vertebrate and yet does not have a heart

—which does full justice to the original (1). So a *generalized conditional,* such as (1), can in full accordance with common usage be construed as affirming a bundle of material conditionals. Taken as a whole, the generalized conditional is a topic for Part II; it lies beyond the present

phase of analysis, which concerns only the compounding of statements explicitly from blocklike components which are self-contained statements in turn.

Another use of 'if-then' which is certainly not to be construed in the fashion of '$p \rightarrow q$' is the *contrafactual conditional;* e.g.:

(4) If Walburg had attended, the measure would have lost.

Whoever affirms a conditional thus in the subjunctive mood is already prepared in advance to maintain also, unconditionally, the falsehood of the antecedent, but still he thinks the conditional adds some information. Surely, then, he does not consider that such a conditional is automatically verified (like '$p \rightarrow q$') simply by the falsity of the antecedent. This kind of conditional is not subject to the earlier remark to the effect that in ordinary usage a conditional is dropped from consideration, as empty and uninteresting, once its antecedent proves false.

The contrafactual conditional is best dissociated from the ordinary conditional in the indicative mood. Whatever the proper analysis of the contrafactual conditional may be, we may be sure in advance that it cannot be truth-functional; for, obviously ordinary usage demands that some contrafactual conditionals with false antecedents and false consequents be true and that other contrafactual conditionals with false antecedents and false consequents be false. Any adequate analysis of the contrafactual conditional must go beyond mere truth values and consider causal connections, or kindred relationships, between matters spoken of in the antecedent of the conditional and matters spoken of in the consequent. It may be wondered, indeed, whether any really coherent theory of the contrafactual conditional of ordinary usage is possible at all, particularly when we imagine trying to adjudicate between such examples as these:

> If Bizet and Verdi had been compatriots, Bizet would have been Italian;
> If Bizet and Verdi had been compatriots, Verdi would have been French.

The problem of contrafactual conditionals is in any case a perplexing one,[7] and it belongs not to pure logic but to the theory of meaning or possibly the philosophy of science. We shall not recur to it here.

[7] See Nelson Goodman, "The problem of counterfactual conditionals."

So the material conditional '$p \rightarrow q$' is put forward not as an analysis of general conditionals such as (1), nor as an analysis of contrafactual conditionals such as (4), but, at most, as an analysis of the ordinary singular conditional in the indicative mood. Even as an analysis of such conditionals the version '$p \rightarrow q$' (or '$-(p\bar{q})$') is sometimes felt to be unnatural, for it directs us to construe a conditional as true no matter how irrelevant its antecedent may be to its consequent, so long as it is not the case that the antecedent is true and the consequent false. The following conditionals, e.g., qualify as true:

(5) If France is in Europe then the sea is salt,
(6) If France is in Australia then the sea is salt,
(7) If France is in Australia then the sea is sweet.

No doubt this result seems strange; but I do not think it would be any less strange to construe (5)–(7) as false. The strangeness is intrinsic rather to the statements (5)–(7) themselves, regardless of their truth or falsity; for it is not usual in practice to form conditionals out of component statements whose truth or falsity is already known unconditionally. The reason this is not usual is readily seen: Why affirm a long statement like (5) or (6) when we are in position to affirm the shorter and stronger statement 'The sea is salt'? And why affirm a long statement like (6) or (7) when we are in position to affirm the shorter and stronger statement 'France is not in Australia'?

In practice, one who affirms 'If p then q' is ordinarily uncertain as to the truth or falsehood individually of 'p' and of 'q' but has some reason merely for disbelieving the combination 'p and not q' as a whole. We say:

If Jones has malaria then he needs quinine,

because we know about malaria but are in doubt both of Jones' ailment and of his need of quinine. Only those conditionals are are worth affirming which follow from some manner of relevance between antecedent and consequent—some law, perhaps, connecting the matters which these two component statements describe. But such connection underlies the useful application of the conditional without needing to participate in its meaning. Such connection underlies the useful application of the conditional even though the meaning of the conditional be understood precisely as '$-(p\bar{q})$'.

The situation is quite similar, indeed, in the case of the connective 'or'. The statement:

France is in Europe or the sea is sweet

is as little worth affirming as (5)–(7) and for the same reason: we can save breath and yet convey more information by affirming simply 'France is in Europe'. In practice one who affirms 'p or q' is ordinarily uncertain as to the truth or falsehood individually of 'p' and of 'q', but believes merely that at least one of the two is true because of a law or some other manner of relevance connecting the matters which the two component statements describe. Yet clearly no meaning need be imputed to 'or' itself beyond the purely truth-functional meaning "not both false".

The question how well '$p \rightarrow q$' conforms to the ordinary indicative 'if-then' is in any case one of linguistic analysis, and of little consequence for our purposes. What is important to note is that '$p \rightarrow q$', the so-called material conditional, is to have precisely the meaning '$-(p\bar{q})$' (or '$\bar{p} \vee q$'); and it will become evident enough, as we proceed, how well adapted this concept is to purposes for which the idiom 'if-then' naturally suggests itself. In particular, as already noted, the material conditional is precisely what is wanted for the individual instances covered by a general conditional of the type (1).

The idiom 'p if and only if q', called the *biconditional,* amounts obviously to the conjunction of two conditionals, 'if p then q' and 'if q then p'. All that has been said regarding the interpretation of the conditional applies *mutatis mutandis* to the biconditional; whatever use 'if-then' may be put to, and whatever meaning it may be conceived to have, a corresponding use and a corresponding meaning must accrue to 'if and only if'. When in particular the conditional is construed as the material conditional '$p \rightarrow q$', the corresponding biconditional is called the *material biconditional* and written '$p \leftrightarrow q$'. Since '$p \leftrightarrow q$' may be regarded simply as an abbreviation of '$(p \rightarrow q)\ (q \rightarrow p)$', or '$-(p\bar{q})\ -(q\bar{p})$', it is evidently false in two and only two cases: in the case where 'p' is true and 'q' false, and in the case where 'q' is true and 'p' false. In other words, a material biconditional is true if the components are alike in truth value (both true or both false), and it is false if the components differ in truth value.

The sign '$\leftrightarrow$', like '$\rightarrow$' and '$\vee$', is dispensable; indeed, we have already seen that '$p \leftrightarrow q$' may be expressed in terms of conjunction and

negation as '$-(p\bar{q})\ -(q\bar{p})$'. But, as will appear in due course, each of these three dispensable signs plays a special part in facilitating the techniques of logic.

HISTORICAL NOTE: The material conditional goes back to Philo of Megara. It was revived in modern logic by Frege (1879) and Peirce (1885). The appropriateness of the material version was vigorously debated in ancient times (see Peirce, 3.441 ff; Łukasiewicz, "Zur Geschichte," p. 116), and has become a current topic of controversy as well. The issue has been clouded, however, by failure to distinguish clearly between the conditional and implication (see Chapter 7).

The sign '$\supset$' was used by Gergonne for the conditional as early as 1816, though not in the material sense, and was revived by Peano and used by Whitehead and Russell in the material sense. Its use has been widespread in subsequent literature, including previous editions of this book. It had the disadvantage, however, of looking like a converse of the inclusion sign '$\subset$' that will emerge in Chapter 20 and is standard usage in set theory. At last I have switched to Hilbert's '$\rightarrow$', which is now widely used and is more suggestive. Conformably I have switched to '$\leftrightarrow$' for the biconditional. My sign for this in previous editions, following Whitehead and Russell, was '$\equiv$', which will be put to a different use in Chapter 20.

EXERCISES

1. It was said in a footnote in Chapter 1 that 'p or q' in the exclusive sense could be written '$p \leftrightarrow \bar{q}$'. Explain why.

2. Translate '$p \leftrightarrow q$' into terms purely of alternation and negation.

4
GROUPING

A conspicuous type of ambiguity in ordinary language is ambiguity of grouping. The statement:

> Rutgers will get the pennant and Hobart will be runner-up if Rzymski is disqualified,

e.g., is hopelessly ambiguous in point of grouping; there is no telling whether Rutgers' getting the pennant is supposed to be contingent upon Rzymski's being disqualified. If so the logical form is '$p \rightarrow qr$', and otherwise it is '$q(p \rightarrow r)$', where 'p' represents 'Rzymski is disqualified', 'q' represents 'Rutgers gets the pennant', and 'r' represents 'Hobart is runner-up'.

In complex statements of ordinary language the intended grouping sometimes has to be guessed, as above, and sometimes has to be inferred from unsystematic cues, as illustrated by the following example:

(1) If the new mail-order campaign does not break the Dripsweet monopoly and restore freedom of competition then Jones will sell his car and mortgage his home.

The words 'if' and 'then' here are helpful in determining the grouping, for they frame the complex antecedent of the conditional just as clearly as if they were parentheses. But they do not show how much text is intended for the consequent of the conditional. Should we stop the consequent of the conditional at the last 'and', or construe it as running clear to the end? The proper answer is evident at a glance; however, let us note explicitly why. The clauses 'Jones will sell his car' and 'Jones will mortgage his home' have been telescoped by omitting the repetition of 'Jones will'; and this affords conclusive evidence that the 'and' here is intended to coordinate just these two clauses, rather than reaching farther back to include a whole conditional as one component of the conjunction. So we know that (1) is to be construed as a conditional, having as antecedent:

the new mail-order campaign does not break the Dripsweet monopoly and restore freedom of competition

and as consequent:

Jones will sell his car and mortgage his home.

But there remains a question of grouping within the antecedent: is the 'not' to govern the whole, or is it to govern just the part preceding 'and'? Obviously the whole. And note that the obviousness of this choice is due to much the same telescoping device as was observed before: the words 'restore freedom of competition' which follow the 'and' must, because of their fragmentary character, be construed as coordinate with 'break the Dripsweet monopoly'. So (1) is a conditional of the form '$-(pq) \rightarrow rs$', where 'p' means 'the new mail-order campaign breaks the Dripsweet monopoly', 'q' means 'the new mail-order campaign restores freedom of competition', 'r' means 'Jones will sell his car', and 's' means 'Jones will mortgage his home'.

We all have an extraordinary finesse at ordinary language; and thus it is that the correctness of the above interpretation of (1) is bound to have been more immediately evident to all of us than the reasons why. But an examination of the reasons affords some notion of the sorts of unsystematic devices whereby ordinary language succeeds in its indications of grouping, such times as it succeeds at all.

We noted the effectiveness of 'if' and 'then' in marking the boundaries of the antecedent of a conditional. In similar fashion 'either' and 'or' may be used to mark the boundaries of the first component of an alternation; and similarly 'both' and 'and' may be used to mark the boundaries of the first component of a conjunction. Thus the ambiguity of:

Jones came and Smith stayed or Robinson left

can, by inserting 'either' at the appropriate point, be resolved in favor of '$pq \vee r$' or '$p(q \vee r)$' at will:

Either Jones came and Smith stayed or Robinson left,
Jones came and either Smith stayed or Robinson left.

Grouping may also be indicated in ordinary language by inserting a vacuous phrase such as 'it is the case that', balanced with another 'that' to

show coordination of clauses. A further device is the insertion of emphatic particles such as 'else' after 'or', or 'also' or 'furthermore' after 'and'; such reinforcement of a connective has the effect of suggesting that it is a major one.

It is evident by now that the artificial notations of logic and mathematics enjoy a great advantage over ordinary language, in their use of parentheses to indicate grouping. Parentheses show groupings unfailingly, and are simple to use. They have the further virtue of allowing complex clauses to be dropped mechanically into place without distortion of clause or of context. This particular virtue has been of incalculable importance; without it mathematics could never have developed beyond a rudimentary stage.

Even so, parentheses can be a nuisance. Unless conventions are adopted for omitting some of them, our longer formulas tend to bristle with them and we find ourselves having to count them in order to pair them off. Actually two conventions for minimizing parentheses have been tacitly in use now for some pages; it is time they were stated. One is this: the connectives '$\vee$', '$\rightarrow$', and '$\leftrightarrow$' are treated as marking a greater break than conjunction. Thus '$pq \vee r$' is understood as having the grouping '$(pq) \vee r$', and not '$p(q \vee r)$'—as is well suggested by the typographical pattern itself. Similarly '$p \vee qr$' means '$p \vee (qr)$', '$pq \rightarrow r$' means '$(pq) \rightarrow r$', etc. The other convention to which we have been tacitly adhering is this: the negation sign is understood as governing as little as possible of what follows it. Thus '$-(pq)r$' means '$(-(pq))r$', not '$-((pq)r)$'; similarly '$-(p \vee q)r$' means '$(-(p \vee q))r$', not '$-((p \vee q)r)$'; and so on.

An auxiliary notation of dots will now be adopted which will have the effect of eliminating *all* parentheses, so far as Part I is concerned, except those directly connected with negation. Perhaps this expedient will seem to reduce parentheses beyond the point of diminishing returns; actually its main value lies in clearing the way for a new influx of parentheses in Part II and beyond.

Dots are reinforcements. They may be thought of as a systematic counterpart of the practice in ordinary language, noted above, of inserting 'else', 'also', etc. To begin with, if we want to convey the meaning '$p(q \vee r)$' and thus create a greater break at the point of conjunction than at the point of alternation, we shall insert a dot at the point of conjunction thus: '$p \,.\, q \vee r$'. For '$(p \vee q)r$' similarly we shall write '$p \vee q \,.\, r$', for '$p(q \rightarrow r)$' we shall write '$p \,.\, q \rightarrow r$', etc.

Next, if at some occurrence of '$\vee$' or '$\rightarrow$' or '$\leftrightarrow$' we want to create a still greater break than is expressed by the dot of conjunction, we shall insert a dot alongside '$\vee$' or '$\rightarrow$' or '$\leftrightarrow$'; thus '$(p \,.\, q \vee r) \leftrightarrow s$' becomes '$p \,.\, q \vee r \,.\leftrightarrow s$'. Just as the undotted '$\vee$' or '$\rightarrow$' or '$\leftrightarrow$' marks a greater break than the undotted conjunction, so the dotted '$\vee$' or '$\rightarrow$' or '$\leftrightarrow$' marks a greater break than the dot of conjunction. The dot which is thus added to reinforce '$\vee$' or '$\rightarrow$' or '$\leftrightarrow$' goes on the side where the reinforcement is needed; thus '$(p \rightarrow q \,.\, r) \vee s$' becomes '$p \rightarrow q \,.\, r \,.\vee s$', but '$p \rightarrow (q \,.\, r \leftrightarrow s)$' becomes '$p \rightarrow .\, q \,.\, r \leftrightarrow s$'. Again '$(p \,.\, q \leftrightarrow r) \vee (p \rightarrow q \,.\, r)$', calling for reinforcement on both sides of the central '$\vee$', becomes '$p \,.\, q \leftrightarrow r \,.\vee.\, p \rightarrow q \,.\, r$'.

When we want to create a still greater break at some point of conjunction than is expressed by a dotted '$\vee$' or '$\rightarrow$' or '$\leftrightarrow$' in the neighborhood, we shall put a double dot '$:$' for the conjunction. When we want to create a still greater break than this at '$\vee$' or '$\rightarrow$' or '$\leftrightarrow$', we shall put a double dot alongside '$\vee$' or '$\rightarrow$' or '$\leftrightarrow$'; and so on to larger groups of dots. What might be written fully in terms of parentheses as:

$$s \vee (p(q \rightarrow r) \leftrightarrow (p \vee q)r)t,$$

e.g., is written with help of dots as follows:

(2) $$s \vee : p \,.\, q \rightarrow r \,.\leftrightarrow.\, p \vee q \,.\, r \,:\, t.$$

In general thus the connectives '$\vee$', '$\rightarrow$', and '$\leftrightarrow$' fare alike. Any group of dots alongside any of these connectives represents a greater break than is represented by the same number of dots standing alone as a sign of conjunction, but a lesser break than is represented by any larger group of dots.

Parentheses will continue to be used to enclose a compound governed by a negation sign; the notations '$-(pq)$', '$-(p \vee q)$', etc. thus persist unchanged. Dots have no power, of course, to transcend parentheses; in '$-(p \vee q \,.\, r)s$', e.g., the dot is powerless to group the 's' with the 'r'.

[8]

We need no parentheses or other indicators of grouping if, with Łukasiewicz, we write each connective before the expressions that it connects instead of writing it between. His notations for 'pq', '$p \vee q$',

[8] Text marked off by the right angles may conveniently be omitted for purposes of a shorter course.

'$p \rightarrow q$', '$p \leftrightarrow q$', and '$\bar{p}$' were 'Kpq', 'Apq', 'Cpq', 'Epq', and 'Np'. The alternation (2) above becomes this in his notation:

(3) AsKEKpC$q$$rKApqrt$.

The grouping is unique and inevitable. The initial 'A' shows that the whole is an alternation whose first component is 's' and whose second component is the rest. The ensuing 'K' shows that the rest is a conjunction. The first component of that conjunction is, in view of the ensuing 'E', a biconditional. Continuing thus, we elicit the intended structure in full.

We have been recognizing conjunctions of not just two components, but of any number; and alternations similarly. In the Łukasiewicz notation this freedom must be curtailed; 'pqr' may be rendered either 'KpKqr', that is, 'p . qr', or 'KKpqr', that is, 'pq . r', these being equivalent; but it must be rendered one way or the other, and not simply 'Kpqr'. Correspondingly for '$p \vee q \vee r$'; it must be rendered 'ApAqr' or 'AApqr'. To permit 'Kpqr' and 'Apqr' would open the way to ambiguity of the following sort: 'KA$pqrs$' could be taken either as 'K(Apqr)s' that is, '$p \vee q \vee r$. s' or as 'K(Apq)rs', that is, '$p \vee q$. rs'. This confusion would matter; the reader can find an assignment of truth values to 'p', 'q', 'r', and 's' that makes '$p \vee q \vee r$. s' true and '$p \vee q$. rs' false.

Łukasiewicz's notation enjoys an arrestingly simple grammar. What strings of letters constitute coherent formulas? 'Cpq' yes; 'ApNKpq' yes; (3) yes; 'pKpNq' no; 'ApKq' no. Here is a simple test, due also to Łukasiewicz. Counting from the beginning of the formula, keep score of the occurrences of statement letters ('p', 'q', etc.) and of two-place connectives ('C', 'A', 'K', 'E'). The occurrences of statement letters outnumber the occurrences of connectives when you reach the end of a proper formula, and not before. I leave the reader to satisfy himself of this rule by experiment.

This notation commands theoretical interest, and it has proved useful in the programming of computers; but dots and parentheses seem more perspicuous.

Students are tempted to tinker with the dot conventions with a view to economy. They note, e.g., that '$p \vee . q \leftrightarrow . r \vee s$' would do in lieu of my '$p \vee : q \leftrightarrow . r \vee s$'; it admits of only the desired interpretation '$p \vee (q \leftrightarrow . r \vee s)$', since '$(p \vee . q \leftrightarrow) r \vee s$' makes no sense. Such economies are ill advised. If an extra dot speeds up the reading, it pays its way. When elegance is what we want, we can get it in full measure by switching to the Łukasiewicz notation.

HISTORICAL NOTE: During the fifteenth, sixteenth, and seventeenth centuries the *vinculum* was commonly used in mathematical writing to indicate grouping. It is an underline or overline. Thus '$p \vee: q \leftrightarrow. r \vee s$' might be rendered '$p \vee \underline{q \leftrightarrow r \vee s}$'. The dots began with Peano and were taken up by Whitehead and Russell and later logicians, subject to varying conventions. The Łukasiewicz notation dates from 1929.[9]

EXERCISES

1. Show how the ambiguous statement:

   John will play or John will sing and Mary will sing

   could be rendered unambiguous, in each of two senses, by telescoping clauses.

2. Indicate and justify the appropriate grouping of:

   If they either drain the swamp and reopen the road or dredge the harbor, they will provide the uplanders with a market and themselves with a bustling trade.

3. Rewrite these using dots:

   $$(p(q \vee r) \rightarrow s) \leftrightarrow (pq \rightarrow s)(pr \rightarrow s),$$
   $$-(p \vee q)(r \vee s) \rightarrow -(p \vee q)s.$$

4. Rewrite this using parentheses:

   $$p \rightarrow. q \vee r \,.\, p \vee qs :\leftrightarrow:. \bar{q}\bar{r} \vee. \bar{p} \,.\, \bar{q} \vee \bar{s} :\rightarrow \bar{p}.$$

5. Rewrite the three above in Łukasiewicz's notation.

[9] See Tarski, *Logic, Semantics, Metamathematics*, p. 39.

5
TRUTH-VALUE ANALYSIS

In Chapter 2 a compound was said to be a truth function of its components when its truth value is determined by those of the components; and it was observed that conjunction and negation constitute an adequate notation for truth functions. In view of this latter circumstance it is natural and convenient hereafter to conceive the notion of truth function in a purely notational way: the *truth functions of* given components are all the compounds constructed from them by means exclusively of conjunction and negation (and the dispensable further connectives '$\vee$', '$\rightarrow$', '$\leftrightarrow$'). Thus '$\bar{p}$' is a truth function of 'p', and '$-(p \vee \bar{r} \mathrel{.\leftrightarrow} pq) \rightarrow r$' is a truth function of 'p', 'q', and 'r'. We also count 'p' itself a truth function of 'p'.

A truth function of letters 'p', 'q', etc., is strictly speaking not a statement, of course, since the letters are themselves not actual statements but mere dummies in place of which any desired statements may be imagined. Hereafter the letters 'p', 'q', etc., and all truth functions of them will be called *schemata* (singular: schema). More specifically they will be called *truth-functional schemata* when it becomes necessary to distinguish them from schemata involving logical devices of other than truth-functional kind. Schemata are logical diagrams of statements; the letters 'p', 'q', etc., by supplanting the component clauses of a statement, serve to blot out all the internal matter which is not germane to the broad outward structures with which our logical study is concerned.

By *interpretation* of the letter 'p' (or 'q', etc.) may be meant specification of an actual statement which is to be imagined in place of the letter. By interpretation of 'p' may also be meant simply specification of a truth value for 'p'. The two senses of 'interpretation' can be used pretty interchangeably because each actual statement S has a specific truth value (known or unknown) and that truth value is all that matters to the truth value of any truth function of S.

A convenient graphic method of imposing interpretations, of the second of the above varieties, is simply to supplant the letters in a schema by the mark '$\top$' for truths and '$\bot$' for falsehoods. Computing then directly with these marks, we can quickly determine what truth value the whole

schema takes on under the imposed interpretations.[10] Thus, suppose our problem is to determine the truth value of the schema '$-(pq \vee \bar{p}\bar{q})$' for the case where '$p$' is interpreted as true and 'q' as false. We simply put 'T' for 'p' and '$\bot$' for 'q' in the schema, getting '$-(\mathsf{T}\bot \vee \bar{\mathsf{T}}\bar{\bot})$'. But, since '$\bar{\mathsf{T}}$' reduces to '$\bot$' and '$\bar{\bot}$' to '$\mathsf{T}$', this becomes '$-(\mathsf{T}\bot \vee \bot\mathsf{T})$'. Further, since a conjunction with false component is false, '$\mathsf{T}\bot$' reduces to '$\bot$' and so does '$\bot\mathsf{T}$'. So the whole is now down to '$-(\bot \vee \bot)$'. But, an alternation of falsehoods being false, '$\bot \vee \bot$' reduces to '$\bot$'; the whole thus becomes '$\bar{\bot}$' or 'T'. This outcome means that our original schema '$-(pq \vee \bar{p}\bar{q})$' comes out true when '$p$' is interpreted as true and 'q' as false.

The process whereby '$-(\mathsf{T}\bot \vee \bar{\mathsf{T}}\bar{\bot})$' was reduced to '$\mathsf{T}$' will be called *resolution*. The simplest of the steps involved in resolution, viz. reduction of '$\bar{\mathsf{T}}$' to '$\bot$' and of '$\bar{\bot}$' to 'T', will always be tacit hereafter; we shall never write '$\bar{\mathsf{T}}$' nor '$\bar{\bot}$', but immediately '$\bot$' and 'T', as if the notation of negation as applied to 'T' and '$\bot$' consisted simply in inverting. The other steps of resolution illustrated in the above example were reduction of '$\mathsf{T}\bot$', '$\bot\mathsf{T}$', and '$\bot \vee \bot$' to '$\bot$'. These steps, and all further ones for which there might be occasion in other examples, may conveniently be codified in the form of nine *rules of resolution:*

(i) *Delete* 'T' *as component of conjunction.* (Thus 'TTT' reduces to 'TT' and thence to 'T'; '$\bot\mathsf{T}$' reduces to '$\bot$'; etc. Reason: a conjunction with a true component is true or false according as the rest of it is true or false.)

(ii) *Delete* '$\bot$' *as component of alternation.* (Thus '$\bot \vee \bot \vee \bot$' reduces to '$\bot \vee \bot$' and thence to '$\bot$'; '$\bot \vee \mathsf{T}$' reduces to 'T'; etc. Reason: an alternation with a false component is true or false according as the rest of it is true or false.)

(iii) *Reduce a conjunction with* '$\bot$' *as component to* '$\bot$'.

(iv) *Reduce an alternation with* 'T' *as component to* 'T'.

(v) *Drop* 'T' *as antecedent of a conditional.* (Reason: a conditional with true antecedent is true or false according as the consequent is true or false.)

(vi) *Reduce a conditional with* '$\bot$' *as antecedent, or* 'T' *as con-*

[10] We need not fumble for a pronunciation of '$\bot$' coordinate with the pronunciation 'tee' of 'T', for the words 'true' and 'false' themselves are short enough to serve conveniently as pronunciations of the two signs. Before deploring my preference of '$\bot$' to the initial 'F' of 'false', note the urgent need of 'F' for other purposes in Parts II–IV.

sequent, to '⊤'. (Thus '⊤ → ⊤', '⊥ → ⊤', and '⊥ → ⊥' reduce to '⊤'.)

(vii) *If a conditional has* '⊥' *as consequent, reduce the whole to the negation of the antecedent.*

(viii) *Drop* '⊤' *as component of a biconditional.* (Thus '⊤ ↔ ⊤' reduces to '⊤', and '⊤ ↔ ⊥' and '⊥ ↔ ⊤' reduce to '⊥'.)

(ix) *Drop* '⊥' *as component of a biconditional and negate the other side.* (Thus '⊥ ↔ ⊥' reduces to '⊤', and '⊤ ↔ ⊥' and '⊥ ↔ ⊤' reduce to '⊥'.)

Set up according to these rules, our original example of resolution amounts to no more than this:

$-(\top\bot \vee \bot\top)$	
$-(\bot \vee \bot)$	(changing '⊤⊥' and '⊥⊤' each to '⊥' by (i) or (iii))
$\top$	(changing '⊥ ∨ ⊥' to '⊥' by (ii))

Turning to a more elaborate example, let us determine the truth value of '$pq \vee \bar{p}\bar{r} .\!\rightarrow\!. q \leftrightarrow r$' for the case where '$p$' and '$q$' are interpreted as false and 'r' as true.

$\bot\bot \vee \top\bot .\!\rightarrow\!. \bot \leftrightarrow \top$	
$\bot\bot \vee \top\bot .\!\rightarrow \bot$	(changing '⊥ ↔ ⊤' to '⊥' by (viii) or (ix))
$-(\bot\bot \vee \top\bot)$	(by (vii))
$-(\bot \vee \bot)$	(by (iii) twice)
$\top$	(changing '⊥ ∨ ⊥' to '⊥' by (ii))

Thus '$pq \vee \bar{p}\bar{r} .\!\rightarrow\!. q \leftrightarrow r$' comes out true when false statements are put for 'p' and 'q' and a true one for 'r'.

Let us feign contact with reality by considering an actual statement of the form '$pq \vee \bar{p}\bar{r} .\!\rightarrow\!. q \leftrightarrow r$':

(1) If either the resident and the deputy resident both resign or the resident neither resigns nor exposes the *chargé d'affaires,* in either case the deputy resident will resign if and only if the resident exposes the *chargé d'affaires.*

What we have found is that (1) comes out true in the case where neither the resident nor the deputy resident resigns and the resident exposes the *chargé d'affaires.*

We have evaluated the schema '$pq \vee \bar{p}\bar{r} .\!\rightarrow\!. q \leftrightarrow r$' for one interpretation: 'p' and 'q' as false and 'r' as true. There remain seven other interpretations that might be considered: 'p', 'q', and 'r' all true, 'p' and 'q' true and 'r' false, 'p' and 'r' true and 'q' false, and so on. The eight cases can be systematically explored, with evaluation of the schema for each case, by the following method. First we put 'T' for 'p', leaving 'q' and 'r' unchanged, and make all possible resolutions by (i)–(ix):

$$\begin{array}{rl} \mathsf{T}q \vee \bot\bar{r} .\!\rightarrow\!. q \leftrightarrow r & \\ q \vee \bot\bar{r} .\!\rightarrow\!. q \leftrightarrow r & \text{(changing '}\mathsf{T}q\text{' to '}q\text{' by (i))} \\ q \vee \bot .\!\rightarrow\!. q \leftrightarrow r & \text{(changing '}\bot\bar{r}\text{' to '}\bot\text{' by (iii))} \\ q \rightarrow. q \leftrightarrow r & \text{(changing '}q \vee \bot\text{' to '}q\text{' by (ii))} \end{array}$$

Then we put 'T' for 'q' in this result and resolve further:

$$\begin{array}{rl} \mathsf{T} \rightarrow. \mathsf{T} \leftrightarrow r & \\ \mathsf{T} \leftrightarrow r & \text{(by (v))} \\ r & \text{(by (viii))} \end{array}$$

We have now found that whenever 'p' and 'q' are both interpreted as true, our original schema resolves to 'r'—hence becomes true or false according as 'r' is true or false. This disposes of two of the eight cases. Next we return to our intermediate result '$q \rightarrow. q \leftrightarrow r$' and put '$\bot$' for '$q$':

$$\begin{array}{rl} \bot \rightarrow. \bot \leftrightarrow r & \\ \mathsf{T} & \text{(by (vi))} \end{array}$$

This shows that our original schema comes out true whenever 'p' is interpreted as true and 'q' as false, regardless of 'r'. This disposes of two more of our eight cases. Now we go all the way back to our original schema and put '$\bot$' for 'p':

$$\begin{array}{rl} \bot q \vee \mathsf{T}\bar{r} .\!\rightarrow\!. q \leftrightarrow r & \\ \bot \vee \mathsf{T}\bar{r} .\!\rightarrow\!. q \leftrightarrow r & \text{(by (iii))} \\ \mathsf{T}\bar{r} \rightarrow. q \leftrightarrow r & \text{(by (ii))} \\ \bar{r} \rightarrow. q \leftrightarrow r & \text{(by (i))} \end{array}$$

Putting 'T' for 'r' here and resolving further, we have:

$$\begin{array}{ll} \bot \rightarrow .\ q \leftrightarrow \top & \\ \top & \text{(by (vi))} \end{array}$$

This shows that our original schema comes out true whenever 'p' is interpreted as false and 'r' as true, regardless of 'q'. Two more cases are disposed of. Finally we go back to '$\bar{r} \rightarrow .\ q \leftrightarrow r$' and put '$\bot$' for '$r$':

$$\begin{array}{ll} \top \rightarrow .\ q \leftrightarrow \bot & \\ q \leftrightarrow \bot & \text{(by (v))} \\ \bar{q} & \text{(by (ix))} \end{array}$$

So whenever 'p' and 'r' are both interpreted as false, our schema resolves to '$\bar{q}$'—hence becomes false or true according as 'q' is interpreted as true or false.

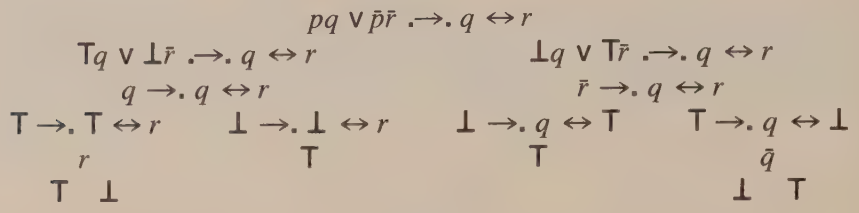
The foregoing analysis might conveniently have been carried out in a single array as follows:

$$\begin{array}{c} pq \vee \bar{p}\bar{r} .\rightarrow .\ q \leftrightarrow r \\ \hline \begin{array}{ll} \top q \vee \bot \bar{r} .\rightarrow .\ q \leftrightarrow r & \bot q \vee \top \bar{r} .\rightarrow .\ q \leftrightarrow r \\ q \vee \bot \bar{r} .\rightarrow .\ q \leftrightarrow r & \bot \vee \top \bar{r} .\rightarrow .\ q \leftrightarrow r \\ q \vee \bot .\rightarrow .\ q \leftrightarrow r & \top \bar{r} \rightarrow .\ q \leftrightarrow r \\ q \rightarrow .\ q \leftrightarrow r & \bar{r} \rightarrow .\ q \leftrightarrow r \\ \hline \top \rightarrow .\ \top \leftrightarrow r \quad \bot \rightarrow .\ \bot \leftrightarrow r & \bot \rightarrow .\ q \leftrightarrow \top \quad \top \rightarrow .\ q \leftrightarrow \bot \\ \top \leftrightarrow r \quad \top & \top \quad q \leftrightarrow \bot \\ r & \bar{q} \\ \top \quad \bot & \bot \quad \top \end{array} \end{array}$$

This is called a *truth-value analysis*. The general method may be summed up as follows. We make a grand dichotomy of cases by putting first '$\top$' and then '$\bot$' for some chosen letter, say 'p'. The expressions thus formed are the respective headings of a bipartite analysis. Then we resolve both expressions, by (i)–(ix), until we end up with '$\top$' or '$\bot$' or some schema. If a schema results, we then proceed to develop, under that schema, a new bipartite analysis with respect to a chosen one of its letters. We continue thus until all end results are single marks—'$\top$' or '$\bot$'. Each end result shows what truth value the original schema will take on when its letters are interpreted according to the marks which have there supplanted them.

Actually all intermeduate steps of resolution are so obvious, and so readily reconstructed at will, that they may hereafter be left to the imagi-

nation. Thus the truth-value analysis above would in future be condensed as follows:


$pq \vee \bar{p}\bar{r} .\rightarrow. q \leftrightarrow r$
$\top q \vee \bot\bar{r} .\rightarrow. q \leftrightarrow r$
$\bot q \vee \top\bar{r} .\rightarrow. q \leftrightarrow r$
$q \rightarrow. q \leftrightarrow r$
$\bar{r} \rightarrow. q \leftrightarrow r$
$\top \rightarrow. \top \leftrightarrow r$
$\bot \rightarrow. \bot \leftrightarrow r$
$\bot \rightarrow. q \leftrightarrow \top$
$\top \rightarrow. q \leftrightarrow \bot$
r
$\top$
$\bar{q}$
$\top \ \bot$
$\bot \ \top$

There is no need always to choose 'p' as the first letter for which to put '$\top$' and '$\bot$'. It is better to choose the letter which has the most repetitions, if repetitions there be, and to adhere to this plan also at each later stage. Thus it was, indeed, that whereas in the second stage on the left side of the above analysis 'q' was chosen for replacement by '$\top$' and '$\bot$', on the other hand in the second stage on the right side 'r' was chosen. This strategy tends to hasten the disappearance of letters, and thus to minimize work.

A method of truth-value analysis that has long been usual in the literature is that of *truth tables*. For the schema '$pq \vee \bar{p}\bar{r} .\rightarrow. q \leftrightarrow r$' last analyzed, the truth table is this:

p	q	r	pq	$\bar{p}$	$\bar{r}$	$\bar{p}\bar{r}$	$pq \vee \bar{p}\bar{r}$	$q \leftrightarrow r$	$pq \vee \bar{p}\bar{r} .\rightarrow. q \leftrightarrow r$
$\top$	$\top$	$\top$	$\top$	$\bot$	$\bot$	$\bot$	$\top$	$\top$	$\top$
$\bot$	$\top$	$\top$	$\bot$	$\top$	$\bot$	$\bot$	$\bot$	$\top$	$\top$
$\top$	$\bot$	$\top$	$\bot$	$\bot$	$\bot$	$\bot$	$\bot$	$\bot$	$\top$
$\bot$	$\bot$	$\top$	$\bot$	$\top$	$\bot$	$\bot$	$\bot$	$\bot$	$\top$
$\top$	$\top$	$\bot$	$\top$	$\bot$	$\top$	$\bot$	$\top$	$\bot$	$\bot$
$\bot$	$\top$	$\bot$	$\bot$	$\top$	$\top$	$\top$	$\top$	$\bot$	$\bot$
$\top$	$\bot$	$\bot$	$\bot$	$\bot$	$\top$	$\bot$	$\bot$	$\top$	$\top$
$\bot$	$\bot$	$\bot$	$\bot$	$\top$	$\top$	$\top$	$\top$	$\top$	$\top$

The first three columns exhaust the assignments of truth values to the letters. The succeeding columns list the truth values of progressively more complex components of the schema for each assignment to the letters. Each of these succeeding columns is derived from earlier columns by reasoning tantamount to what I have called resolution.

The truth table can be constructed more compactly thus:

p	q	$\vee$	$\bar{p}$	$\bar{r}$	$.\rightarrow.$	$q \leftrightarrow r$
T	T	T	⊥	⊥	T	T
⊥	⊥	T	⊥	⊥	T	T
T	⊥	T	⊥	⊥	T	⊥
⊥	⊥	⊥	T	T	T	⊥
T	T	T	⊥	T	⊥	T
⊥	T	T	T	T	⊥	T
T	⊥	T	⊥	T	T	⊥
⊥	T	T	T	T	T	⊥

We begin this construction by writing out the schema and inscribing the appropriate columns under its single letters and their negations. Then we derive the column for 'pq', building it under the middle of that conjunction. Similarly for '$\bar{p}\bar{r}$' and '$q \leftrightarrow r$'. Then we derive the column for the alternation, building it under the '$\vee$'; and finally we build the column for the whole schema under its main connective, '$\rightarrow$'.

But the branching notation of truth-value analysis that we just previously arrived at is, we see, briefer still than this compacted truth table. The saving increases markedly, moreover, when the number of distinct letters is increased, as the reader can verify by experiment. For four letters, after all, the truth table runs to sixteen lines; for five, thirty-two. The saving increases further when, as in the next chapter, the analysis specifically seeks consistency or validity; for these questions are answered by truth tables only when the tables are nearly finished, whereas under our branching procedure of truth-value analysis they often are answered early in the game.

HISTORICAL NOTE: The pattern of reasoning that the truth table tabulates was Frege's, Peirce's, and Schröder's by 1880. The tables have been prominent in the literature since 1920 (Łukasiewicz, Post, Wittgenstein). The compact style last displayed above is from my *Mathematical Logic*, 1940.

EXERCISES

1. Suppose they drain the swamp but neither reopen the road nor dredge the harbor nor provide the uplanders with a market; and suppose nevertheless they do provide themselves with a bustling trade. Determine, under these circumstances, the truth value of the statement in Exercise 2 of the preceding chapter. Method: represent the components as 'p', 'q', 'r', 's', 't'; put 'T' and '$\perp$' appropriately for the letters; resolve.

2. Suppose they neither drain the swamp nor reopen the road but that they both dredge the harbor and provide the uplanders with a market, and still do not provide themselves with a bustling trade. What then is the truth value of the statement in Exercise 2 of the preceding chapter?

3. Make a truth-value analysis of each of the schemata:

$$p \rightarrow pq, \qquad p \rightarrow .\, p \vee q, \qquad p \rightarrow .\, p \rightarrow q, \qquad p \rightarrow .\, p \leftrightarrow q.$$

   Present your work in full, showing intermediate steps of resolution; afterward circle those intermediate lines for omission, to show how the work would look in the condensed style.

4. For comparison perform two truth-value analyses of '$p \rightarrow q \,.\, q \rightarrow r$', first following and then flouting the strategy of choosing the most frequent letter.

6
CONSISTENCY AND VALIDITY

A truth-functional schema is called *consistent* if it comes out true under some interpretation of its letters; otherwise *inconsistent*. A truth-functional schema is called *valid* if it comes out true under every interpretation of its letters. The schema '$p\bar{q}$', for example, is consistent, for

it comes out true when 'p' is interpreted as true and 'q' as false; but it is not valid, since there are other interpretations of 'p' and 'q' which make it come out false.

The way to test a truth-functional schema for validity and consistency is obvious: we carry out a truth-value analysis and see whether we get '$\top$' in every case (showing validity) or '$\bot$' in every case (showing inconsistency) or neither. Two examples of validity and inconsistency are respectively '$p \rightarrow p$' and '$p\bar{p}$':

$$\begin{array}{cccc} & p \rightarrow p & & p\bar{p} \\ \top \rightarrow \top & & \bot \rightarrow \bot & \top\bot \qquad \bot\top \\ \top & & \top & \bot \qquad \bot \end{array}$$

Valid schemata were already exploited at one point in the argument of Chapter 2, where '$-(p\bar{p}qrs)$' was cited.

$$\begin{array}{ccc} & -(p\bar{p}qrs) & \\ -(\top\bot qrs) & & -(\bot\top qrs) \\ \top & & \top \end{array}$$

In general, obviously, a schema is valid if and only if its negation is inconsistent, and a schema is inconsistent if and only if its negation is valid. Thus the negations '$-(p\bar{p})$' and '$-(p\bar{p}qrs)$' of the inconsistent schemata '$p\bar{p}$' and '$p\bar{p}qrs$' are valid, and the negation '$-(p \rightarrow p)$' of the valid schema '$p \rightarrow p$' is inconsistent.

A test of validity may be stopped short, with negative outcome, as soon as we come to a case yielding '$\bot$'; and a test of consistency may be stopped, with affirmative outcome, as soon as we come to a case yielding '$\top$'. Thus the analysis of '$pq \vee \bar{p}\bar{r} .\rightarrow. q \leftrightarrow r$' in Chapter 5 might, if we had been interested only in consistency and validity, have been discontinued in this fragmentary state:

$$\begin{array}{l} \qquad\qquad\qquad\qquad pq \vee \bar{p}\bar{r} .\rightarrow. q \leftrightarrow r \\ \qquad \top q \vee \bot\bar{r} .\rightarrow. q \leftrightarrow r \\ \qquad\quad q \rightarrow. q \leftrightarrow r \\ \top \rightarrow. \top \leftrightarrow r \\ \qquad r \\ \quad \top \quad \bot \end{array}$$

This much already suffices to show both that '$pq \vee \bar{p}\bar{r} .\rightarrow. q \leftrightarrow r$' is consistent and that it is not valid.

'Validity' is not to be thought of as a term of praise. When a schema is valid, any statement whose form that schemata depicts is bound to be, in some sense, trivial. It will be trivial in the sense that it conveys no real information regarding the subject matter whereof its component clauses speak. The statement:

(1) If the Bruins win then the Bruins win,

whose form is depicted in the valid schema '$p \rightarrow p$', gives us no information about the outcome of the game; indeed, any other clause, on any other subject matter, could be used here in place of 'the Bruins win' with as much and as little effect. Valid schemata are important not as an end but as a means. We shall see in another page or two that simple cases of validity afford shortcuts in the truth-value analysis of other schemata; and we shall see in the next chapter that the determination of certain complex cases of validity is tantamount to determining relations of equivalence and implication between other schemata.

Though the statements illustrative of valid schemata are always trivial in the sense noted above, they are not always trivial in the sense of being, like (1), recognizable on sight. Valid schemata may run to any length and any degree of complexity; and some even of moderate length cannot be recognized as valid without substantial computation. The same is true of consistency. Below is a schema of quite moderate complexity which, though by no means recognizable as valid on sight, is found to be valid by truth-value analysis.

$$pq \vee p\bar{r} \vee \bar{p}r \vee \bar{p}s \vee \bar{q}r \vee \bar{r}\bar{s}$$

$$\mathsf{T}q \vee \mathsf{T}\bar{r} \vee \bot r \vee \bot s \vee \bar{q}r \vee \bar{r}\bar{s} \qquad \bot q \vee \bot\bar{r} \vee \mathsf{T}r \vee \mathsf{T}s \vee \bar{q}r \vee \bar{r}\bar{s}$$

$$q \vee \bar{r} \vee \bar{q}r \vee \bar{r}\bar{s} \qquad r \vee s \vee \bar{q}r \vee \bar{r}\bar{s}$$

$$q \vee \bot \vee \bar{q}\mathsf{T} \vee \bot\bar{s} \qquad q \vee \mathsf{T} \vee \bar{q}\bot \vee \mathsf{T}\bar{s} \qquad \mathsf{T} \vee s \vee \bar{q}\mathsf{T} \vee \bot\bar{s} \qquad \bot \vee s \vee \bar{q}\bot \vee \mathsf{T}\bar{s}$$

$$q \vee \bar{q} \qquad \mathsf{T} \qquad \mathsf{T} \qquad s \vee \bar{s}$$

$$\mathsf{T} \vee \bot \quad \bot \vee \mathsf{T} \qquad\qquad \mathsf{T} \vee \bot \quad \bot \vee \mathsf{T}$$

$$\mathsf{T} \quad \mathsf{T} \qquad\qquad \mathsf{T} \quad \mathsf{T}$$

The test shows validity, but unimplemented inspection would have availed little.

Taking advantage of the really evident cases of validity and inconsistency, however, we may speed up our truth-value analyses hereafter. Schemata like '$q \vee \bar{q}$' and '$s \vee \bar{s}$', which emerged in the course of the analysis above, are now known to be valid, and hence to reduce to 'T' in

all cases; hereafter, therefore, we may as well agree to reduce any such result directly to '⊤', without further ceremony. Thus, in place of the configurations:

$$\begin{array}{ccc} & q \vee \bar{q} & \\ \top \vee \bot & & \bot \vee \top \\ \top & & \top \end{array} \qquad \begin{array}{ccc} & s \vee \bar{s} & \\ \top \vee \bot & & \bot \vee \top \\ \top & & \top \end{array}$$

which appeared in the lower corners of the analysis above, we shall in future write simply:

$$\begin{array}{c} q \vee \bar{q} \\ \top \end{array} \qquad \begin{array}{c} s \vee \bar{s} \\ \top \end{array}$$

In general, any such patently valid schema may be reduced immediately to '⊤' whenever it occurs in a truth-value analysis, whether in isolation or as component of a longer formula. Similarly any patently inconsistent schema such as '$p\bar{p}$', '$p\bar{p}qrs$', '$\bar{q}pqr$', etc. may be reduced immediately to '⊥' whenever it turns up in an analysis.

With these shortcuts in mind, let us analyze a really complex schema:

$$\begin{array}{c} p \vee q \;.\; p \vee \bar{q} \;.\vee\; \bar{p}q \;:\leftrightarrow q \;.:\rightarrow.\; pr \vee p\bar{r} \\ \top \vee q \;.\; \top \vee \bar{q} \;.\vee\; \bot q \;:\leftrightarrow q \;.:\rightarrow.\; \top r \vee \top \bar{r} \\ q \rightarrow. \; r \vee \bar{r} \\ q \rightarrow \top \\ \top \end{array} \qquad \begin{array}{c} \\ \bot \vee q \;.\; \bot \vee \bar{q} \;.\vee\; \top q \;:\leftrightarrow q \;.:\rightarrow.\; \bot r \vee \bot \bar{r} \\ -(q\bar{q} \vee q \;.\leftrightarrow q) \\ -(\bot \vee q \;.\leftrightarrow q) \\ -(q \leftrightarrow q) \\ \bot \end{array}$$

The reduction of the left side to '$q \rightarrow .\; r \vee \bar{r}$' proceeded by various steps of resolution, as usual. But in the next step, using our new shortcut, we put '⊤' for the patently valid '$r \vee \bar{r}$' and got '$q \rightarrow \top$', which then resolved into '⊤'. On the right-hand side of the analysis, the reduction to '$-(q\bar{q} \vee q \;.\leftrightarrow q)$' proceeded by resolution as usual; then, using our new shortcut, we put '⊥' for the patently inconsistent '$q\bar{q}$', and got '$-(\bot \vee q \;.\leftrightarrow q)$', which resolved in turn into '$-(q \leftrightarrow q)$'. Here, using our new shortcut again, we put '⊤' for the patently valid '$q \leftrightarrow q$' and conclude our work.

Where to draw the line between what is patently valid or inconsistent and what is not patently so is quite arbitrary. The '$-(q\bar{q} \vee q \;.\leftrightarrow q)$' in the right-hand part of the analysis above is itself an inconsistent schema, and so might have been supplanted immediately by '⊥' if its inconsistency

had been felt to be sufficiently obvious. Similarly the '$q \rightarrow . r \vee \bar{r}$' in the left-hand part of the analysis might have been supplanted directly by '$\top$'. For uniformity of classroom work, we might limit the category of "patently inconsistent" schemata to these two kinds: (a) conjunctions such as '$\bar{q}pqr$', '$p \vee q \,.\, r \,.\, -(p \vee q)$', etc., in which some part appears both plain and negated as component of the conjunction, and (b) biconditionals like '$p \leftrightarrow \bar{p}$' or '$-(qr) \leftrightarrow qr$'. But beware of '$p \rightarrow \bar{p}$'; it is consistent.

We might limit the category of "patently valid" schemata to these two kinds: (a) alternations such as '$\bar{q} \vee p \vee q \vee r$', '$pq \vee r \vee -(pq)$', etc., in which some part appears both plain and negated as component of the alternation; (b) conditionals or biconditionals whose two sides are alike, e.g., '$q \leftrightarrow q$', '$qr \leftrightarrow qr$', '$p \vee q \mathrel{.\!\!\rightarrow} . p \vee q$'.

It is only to such schemata, then that our shortcut is to be applied; these, and only these, will be reduced appropriately to '$\bot$' or '$\top$' on sight.

From the validity of a schema we may infer, without separate test, the validity of any schema which is formed from it by *substitution*. From the validity, e.g., of '$p \vee \bar{p}$' we may infer the validity of the schema '$qr \vee -(qr)$' which is formed from '$p \vee \bar{p}$' by substituting 'qr' for 'p'. This is apparent from the definition of validity. Validity of '$p \vee \bar{p}$' means that '$p \vee \bar{p}$' is bound to come out true no matter what statement be put for 'p'; so it follows, as a special case, that '$qr \vee -(qr)$' will come out true no matter what statement 'qr' be made to represent—hence no matter what statements be put for 'q' and 'r'. *Substitution of schemata for letters preserves validity*. But it is clearly essential that 'substitution for a letter' be construed as meaning uniform substitution for every occurrence of the letter. From the validity of '$p \vee \bar{p}$', e.g., we are not entitled to infer validity of '$qr \vee \bar{p}$' nor of '$qr \vee -(qs)$'. It is permissible to put the same or different schemata for different letters, but we must always put the same schema for recurrences of the same letter.

Since inconsistency of a schema is simply validity of its negation, we may conclude further that *substitution of schemata for letters preserves inconsistency*. But note on the other hand that substitution cannot be depended upon to preserve consistency. The mere fact that the more general schema has *some* true instances (which is what consistency means) gives us no reason to suppose that the special case will share any of the true instances. The schema '$p \vee pq$', e.g., is consistent (as may be verified by truth-value analysis), but substitution of '$r\bar{r}$' for 'p' therein yields an inconsistent schema '$r\bar{r} \vee r\bar{r}q$'. Similarly, substitution for a letter in a nonvalid schema cannot be depended upon to yield a nonvalid schema; it may yield a valid or nonvalid one.

EXERCISES

1. Test each of these for validity by truth-value analysis, exploiting the new shortcut regarding patently valid and patently inconsistent clauses:

$$p \rightarrow q \;.\vee.\; q \rightarrow p, \qquad p \vee qr \vee. \bar{p} \;.\; \bar{q} \vee \bar{r},$$
$$p \leftrightarrow q \;.\vee.\; p \leftrightarrow \bar{q}, \qquad p \leftrightarrow q \;.\vee.\; q \leftrightarrow r \;.\vee.\; p \leftrightarrow r.$$

2. In each of the four schemata above, substitute '$p \vee q$' for 'p'. This is chiefly an exercise in adjusting dots to preserve proper grouping.

3. "If a schema is consistent but not valid, then by one set of substitutions we can get a valid schema from it, and by another set of substitutions an inconsistent schema." Is this true? Justify your answer.

4. By negating a consistent schema can you get a valid one? a consistent one? an inconsistent one? Illustrate your affirmative answers.

7
IMPLICATION

The most conspicuous purpose of logic, in its applications to science and everyday discourse, is the justification and criticism of inference. Logic is largely concerned with devising techniques for showing that a given statement does, or does not, "follow logically" from another. The statement 'No dropped freshman is eligible for the Bowdoin Prize', e.g., follows logically from 'No freshman is eligible for the Bowdoin or Bechtel Prize'; and the statement 'Cassius is not both lean and hungry' follows logically from 'Cassius is not hungry'. Now the first of these two examples lies beyond the scope of the truth-functional part of logic with which we are concerned in Part I, but the second example can already be treated here.

From the point of view of logical theory, the fact that the statement 'Cassius is not both lean and hungry' follows from 'Cassius is not hungry'

is conveniently analyzed into these two circumstances: (a) the two statements have the respective logical forms '$-(pq)$' and '$\bar{q}$' (with 'Cassius is lean' and 'Cassius is hungry' supplanting 'p' and 'q'); and (b) there are no two statements which, put respectively for 'p' and 'q', make '$\bar{q}$' true and '$-(pq)$' false. Circumstance (b) will hereafter be phrased in this way: '$\bar{q}$' *implies* '$-(pq)$'. In general, one truth-functional schema is said to *imply* another if there is no way of so interpreting the letters as to make the first schema true and the second false.

Whether a truth-functional schema S_1 implies another, S_2, can be decided always by taking S_1 as antecedent and S_2 as consequent of a conditional, and testing the conditional for validity. For, according to our definition, S_1 implies S_2 if and only if no interpretation makes S_1 true and S_2 false, hence if and only if no interpretation falsifies the material conditional whose antecedent is S_1 and whose consequent is S_2. In a word, *implication is validity of the conditional*. To determine that '$\bar{q}$' implies '$-(pq)$', e.g., we check the validity of the corresponding conditional:

$$\bar{q} \rightarrow -(pq)$$

$$\begin{array}{c} \bot \rightarrow -(p\mathsf{T}) \\ \mathsf{T} \end{array} \qquad\qquad \begin{array}{c} \mathsf{T} \rightarrow -(p\bot) \\ \mathsf{T} \end{array}$$

Next let us note an example which turns out negatively. That '$p \vee q$' does *not* imply 'pq' is found thus:

$$p \vee q \mathrel{.\!\!\rightarrow} pq$$

$$\begin{array}{c} \mathsf{T} \vee q \mathrel{.\!\!\rightarrow} \mathsf{T}q \\ q \\ \mathsf{T} \quad \bot \end{array}$$

Once having come out with a '$\bot$', we discontinue our test in the knowledge that '$p \vee q \mathrel{.\!\!\rightarrow} pq$' is not valid; i.e., that '$p \vee q$' does not imply 'pq'. This result does not mean that '$p \vee q \mathrel{.\!\!\rightarrow} pq$' does not come out true under *some* interpretations of 'p' and 'q', nor does it mean that '$p \vee q$' and 'pq' themselves do not come out simultaneously true under some interpretations of 'p' and 'q'. The failure of implication means merely that *some* interpretations which make '$p \vee q$' true make 'pq' false; or, what comes to the same thing, that some interpretations make '$p \vee q \mathrel{.\!\!\rightarrow} pq$' false.

By reflecting briefly on our methods of testing for implication, validity, and inconsistency, one sees that these four general laws hold:

(i) Any schema implies itself.

(ii) If one schema implies a second and the second a third then the first implies the third.

(iii) An inconsistent schema implies every schema and is implied by inconsistent ones only.

(iv) A valid schema is implied by every schema and implies valid ones only.

We are told in the second half of (iv) that implication, like substitution, transmits validity. Substitution does so, we saw, because all interpretations of the second schema are interpretations of the first. Implication does so for a rather opposite reason: because all interpretations that make the *first* schema *true* make the second true.

An easy familiarity with simple cases of implication between truth-functional schemata will be found to facilitate construction of proofs at even as advanced a level of logic as Chapters 38 ff. At that stage it will not be enough to be able to answer raised questions of implication, which we can do by truth-value analysis as above; we must also be able to raise the questions. We must be able to think up schemata which imply or are implied by a given schema and promise well as links in a proposed chain of argument. Such products of imagination can be checked mechanically by truth-value analysis, but thinking them up is an unmechanical activity. Facility in it depends on grasping the sense of simple schemata clearly enough to be able, given a schema, to conjure up quite an array of fairly simple variants which imply or are implied by it. Given '$p \vee q$', it should occur to us immediately that 'p' and 'q' and 'pq' and '$\bar{p} \rightarrow q$' imply it and that '$p \vee q \vee r$' and '$\bar{p} \rightarrow q$' are implied by it. Given '$p \rightarrow q$', it should occur to us immediately that each of:

$$\bar{p}, \quad q, \quad qr, \quad \bar{p} \vee q, \quad \bar{q} \rightarrow \bar{p}, \quad p \rightarrow qr, \quad p \vee r .\!\rightarrow q$$

implies it and that each of:

$$\bar{p} \vee q, \quad \bar{q} \rightarrow \bar{p}, \quad p \rightarrow .\, q \vee r, \quad p \rightarrow q \,.\vee r$$

is implied by it. Such flashes need not be highly accurate, for we can check each hunch afterward by truth-value analysis. What is important is that they be prolific, and accurate enough to spare excessive lost motion.

No doubt repertoire is an aid to virtuosity in contriving implications, but understanding is the principal thing. When simple schemata are

sufficiently transparent to us, we can see through them by the light of pure reason to other schemata which must come out true if these do, or which can not come out true unless these do. It is well to reflect upon the above examples and succeeding ones until it becomes obvious from the sheer meanings of signs that the implications must hold.

Readiness with implications is aided also, no doubt, by ease of checking. Accordingly a quick implication test called the *fell swoop* will now be explained which, though not general, works for an important range of simple cases.

Some schemata are visibly verifiable by one and only one interpretation of their letters. E.g., '$p\bar{q}$' comes out true when and only when '$\top$' is put for 'p' and '$\bot$' for 'q'. Now when S is such a schema, the question whether S implies a schema S' can be settled simply by supplanting 'p', 'q', etc., in S' by the values which make for truth of S, and resolving. If we come out with '$\top$' or a valid schema, then S implies S'; otherwise not. E.g., to determine that '$p\bar{q}$' implies '$p \to q \mathrel{.\!\to} r$' we put '$\top$' for '$p$' and '$\bot$' for '$q$' in '$p \to q \mathrel{.\!\to} r$' and resolve the result '$\top \to \bot \mathrel{.\!\to} r$', getting '$\top$'.

In particular a fell swoop will settle any question of implication on the part of 'p' or '$\bar{p}$'. To find that 'p' implies '$q \to p$' we put '$\top$' for 'p' in '$q \to p$' and resolve the result '$q \to \top$' to '$\top$'. To find that 'p' implies '$p \to q \mathrel{.\!\to} q$' we put '$\top$' for '$p$' in '$p \to q \mathrel{.\!\to} q$' and resolve the result '$\top \to q \mathrel{.\!\to} q$', coming out with the valid schema '$q \to q$'. To find that '$\bar{q}$' implies '$-(pq)$', which was the example of Cassius, we could have simply put '$\bot$' for 'q' in '$-(pq)$' and resolved the result '$-(p\bot)$' to '$\top$'.

Some schemata, on the other hand, are visibly falsifiable by one and only one interpretation of their letters. E.g., '$-(pr)$' comes out false when and only when '$\top$' is put for 'p' and 'r'; '$p \to r$' comes out false when and only when '$\top$' is put for 'p' and '$\bot$' for 'r'; '$p \vee r$' comes out false when and only when '$\bot$' is put for 'p' and 'r'; '$pr \to s$' comes out false when and only when '$\top$' is put for 'p' and 'r' and '$\bot$' for 's'; and '$p \to .\, r \vee s$' comes out false when and only when '$\top$' is put for 'p' and '$\bot$' for 'r' and 's'. Now when S' is a schema thus falsifiable by one and only one interpretation, the question whether a schema S implies S' can be settled simply by supplanting 'p', 'q', etc. in S by the values which make for falsity of S', and resolving. If we come out with '$\bot$' or an inconsistent schema, then S implies S'; otherwise not. For, the implication can fail only through truth of S where S' is false.

E.g., to find that '$p \to q \,.\, q \to r$' implies '$p \to r$' we put '$\top$' for 'p' and '$\bot$' for 'r' in '$p \to q \,.\, q \to r$' and resolve the re-

sult '$\top \rightarrow q \,.\, q \rightarrow \bot$', getting the inconsistent schema '$q\bar{q}$'. To find that '$p \vee q \,.\, q \rightarrow r$' implies '$p \vee r$' we put '$\bot$' for '$p$' and '$r$' in '$p \vee q \,.\, q \rightarrow r$' and resolve, getting '$q\bar{q}$' again. To find that '$p \rightarrow q \,.\, qr \rightarrow s$' implies '$pr \rightarrow s$' we put '$\top$' for '$p$' and '$r$' and '$\bot$' for '$s$' in '$p \rightarrow q \,.\, qr \rightarrow s$' and resolve.

In particular this backward variety of the fell swoop is convenient when we want to know whether a schema S implies 'p', or '$\bar{p}$'. To find that '$pq \vee p\bar{q}$' implies 'p' we put '$\bot$' for 'p' in '$pq \vee p\bar{q}$' and resolve the result '$\bot q \vee \bot\bar{q}$', getting '$\bot$'. To find that '$p \vee q \,.\, p \vee \bar{q}$' implies '$p$' we put '$\bot$' for '$p$' in '$p \vee q \,.\, p \vee \bar{q}$' and resolve the result '$\bot \vee q \,.\, \bot \vee \bar{q}$', getting the inconsistent schema '$q\bar{q}$'.

Fell swoops are possible only where the schema which is to do the implying clearly comes out true under one and only one interpretation, or else the schema which is to be implied comes out false under one and only one interpretation. The general test of implication, applicable in every case, is truth-value analysis of the conditional; the *full sweep* as opposed to the fell swoop.

Implication may be made to relate statements as well as schemata. When one schema implies another, and a pair of statements are obtained from the schemata by interpretation, we may say by extension that the one statement implies the other. Thus, besides saying that '$\bar{q}$' implies '$-(pq)$', we may make interpretations and say that 'Cassius is not hungry' implies 'Cassius is not lean and hungry'. But it is well here to say more explicitly that the one statement implies the other *truth-functionally*, adding the adverb as a reminder that the schemata which brought the two statements into an implication relationship were truth-functional schemata rather than schemata of kinds which have yet to be taken up in Part II and beyond. Truth-functional implication is, in other words, the relation which one statement bears to another when the second follows from the first by logical considerations within the scope of the logic of truth functions. The terms 'truth-functionally valid' and 'truth-functionally inconsistent' may be applied to statements in similar fashion.

Implication, as we have seen, is intimately related to the conditional. Implication holds when and only when the conditional is valid. This important connection has engendered a tendency among writers on logic to adopt 'implies', confusingly, as a reading of the conditional sign '$\rightarrow$' itself. Then, since '$p \rightarrow q$' has been explained as coming out true whenever 'p' is interpreted as false or 'q' as true, it is concluded with an air of paradox that every falsehood implies every statement and that every

truth is implied by every statement. It is not perceived that '→' is at best an approximation to 'if-then', not to 'implies'.

In order fully to appreciate the distinction which I intend between '→', or 'if-then', and 'implies', it is necessary to become clearly aware of the difference between use and mention. When we say that Cambridge adjoins Boston we mention Cambridge and Boston, but use the names 'Cambridge' and 'Boston'; we write the verb 'adjoins' not between Cambridge and Boston, but between their names. When the mentioned objects are cities, as here, use and mention are unlikely to be confused. But the same distinction holds when the mentioned objects are themselves linguistic expressions. When we write:

The fifth word of ''The Raven'' rhymes with the eleventh

we mention the words 'dreary' and 'weary', but what we use are names of them. We write 'rhymes with' not between the rhyming words but between their names. We may also write:

'dreary' rhymes with 'weary',

but here again we are using names of the rhyming words in question—the names being in this case formed by adding single quotation marks. It would be not merely untrue but ungrammatical and meaningless to write:

Dreary rhymes with weary.

Now when we say that one statement or schema implies another, similarly, we are not to write 'implies' between the statements or schemata concerned, but between their names. In this way we mention the schemata or statements, we talk *about* them, but use their names. These names are usually formed by adding single quotation marks.[11] Validity and consistency are in this respect on the same footing with implication; we say that a schema or statement is valid or consistent by appending 'is valid' or 'is consistent' not to the schema or statement in question but to a name of it.

When on the other hand we compound a statement or schema from two others by means of 'if-then', or '→', we use the statements or schemata themselves and not their names. Here we do not *mention* the statements or

[11] When the expression to be named is displayed in an isolated line or lines, I make a colon do the work of single quotation marks; see above.

schemata. There is no reference to them; they merely occur as parts of a longer statement or schema. The conditional:

If Cassius is not hungry then he is not lean and hungry

mentions Cassius, and says something quite trivial about him, but it mentions no statements at all. The situation here is the same as with conjunction, alternation, and negation.

We have made a point of handling 'if-then' truth-functionally. Among our topics of logical analysis, indeed, no place has been made for non-truth-functional ways of compounding statements. But the fact remains that implication, as a relation between statements, imputes intimate structural connections; it involves far more than the mere truth values of the two statements. This fact conflicts in no way with a strict adherence to truth-functional ways of *compounding* statements and schemata, insofar as statements or schemata are to be compounded at all. The verbs 'implies', 'is longer than', 'is clearer than', and 'rhymes with' are all on a par so far as the present contrasts are concerned: they connect, not statements to form compound statements, but names of statements to form statements about statements.

HISTORICAL NOTE: The distinction that was stressed just now was woefully neglected by Whitehead and Russell, who accorded '$p \rightarrow q$' the readings 'if p then q' and 'p implies q' indifferently. The old controversy over the material conditional (Chapter 3) was, in consequence, aggravated. The truth function '$-(p\bar{q})$' meets some opposition as a rendering of 'if-then'; it meets more, and rightly, as a rendering of 'implies'. Certainly implication must be preserved as a strong relation, dependent upon the structure of the related statements and not just the truth values.

The habit of pronouncing '$\rightarrow$' as 'implies' still persists, and is to be deplored. Partly it is encouraged by the trivial circumstance that 'if' breaks the word order, while 'implies' falls into the position of '$\rightarrow$'. Anyone thus tempted should observe that '$\rightarrow$' can be read without change of position as 'only if'; see the next chapter.

EXERCISES

1. Determine by truth-value analysis whether '$p \leftrightarrow . q \leftrightarrow r$' implies '$r \leftrightarrow . q \rightarrow p$' or vice versa.

2. Do the same for these:

$$p \leftrightarrow qr \mathrel{.\!\rightarrow} pq, \qquad p \rightarrow qr \mathrel{.\!\leftrightarrow} pq.$$

3. Determine which of the four schemata:

$$p \mathrel{.} p \rightarrow q, \qquad \bar{q} \mathrel{.} p \rightarrow q, \qquad p \mathrel{.} \bar{p} \rightarrow q, \qquad p \rightarrow q\bar{q}$$

imply 'q' and which imply '$\bar{p}$'. This means eight fell swoops.

4. Determine what implications hold between these:

$$p \rightarrow q, \qquad p \vee q \mathrel{.\!\rightarrow} r, \qquad p \rightarrow\!\!. \; q \vee r.$$

Note that '$p \rightarrow\!\!. \; q \vee r$', like '$p \rightarrow q$', becomes false under just one interpretation of its letters.

5. Determine what implications hold between these:

$$p\bar{q}, \qquad \bar{p} \leftrightarrow q, \qquad p \vee q.$$

6. Find as many schemata as you can, containing one occurrence each of 'p' and 'q' and no further letters, such that each implies '$\bar{p}$'. Also find as many as you can which are implied by '$\bar{p}$'.

7. Determine whether either of these statements implies the other:

> The company is responsible if and only if the unit was an Interplex and installed since January.
>
> If the unit was an Interplex, then it was installed since January and the company is responsible; and if the unit was not an Interplex, then it was not installed since January and the company is not responsible.

Method: Obtain schemata representing the logical forms of these statements by using 'p', 'q', and 'r' for the component statements; then test the schemata for implication. Be sure to use 'p' for one and the same component throughout both compounds and similarly for 'q' and 'r'. Be sure also to keep the proper groupings.

8
WORDS INTO SYMBOLS

Logical inference leads from *premises*—statements assumed or believed for whatever reason—to *conclusions* which can be shown on purely logical grounds to be true if the premises are true. Techniques to this end are a primary business of logic, and have already begun to occupy our attention. But whereas the connection between premises and conclusions is thus grounded in logic, ordinarily the premises and conclusions themselves are not; and herein precisely lies the *application* of logic to fields other than itself.

The premises and conclusions may treat of any topics and are couched, to begin with, in ordinary language rather than in the technical ideography of modern logic. It is as an aid to establishing implications that we then proceed to mutilate and distort the statements, introducing schematic letters in order to bring out relevant skeletal structures, and translating varied words into a few fixed symbols such as '$\rightarrow$' and '$\vee$' in order to gain a manageable economy of structural elements. The task of thus suitably paraphrasing a statement and isolating the relevant structure is just as essential to the application of logic as is the test or proof of implication for which that preliminary task prepares the way.

An example of how such paraphrasing reduces varied idioms to uniformity has already been noted in the notation of negation (see Chapter 1). The notation of conjunction has a similar effect; for in ordinary language conjunction is expressed not only by 'and' but also by 'but', by 'although', by unspoken punctuation, and in various other ways. Consideration of 'but' and 'although' is instructive, for it brings out a distinction between what may be called the logical and the rhetorical aspects of language. We are likely to say:

Jones is here but Smith is away,

rather than:

Jones is here and Smith is away,

because of the contrast between being here and being away; or, if the contrast between 'Jones is here' and 'Smith is away' attains such proportions as to cause surprise, as it might, e.g., if Jones is not in the habit of coming except to see Smith, we are likely to say:

Jones is here although Smith is away.

But the circumstances which render the compound true are always the same, viz., joint truth of the two components, regardless of whether 'and', 'but', or 'although' is used. Use of one of these words rather than another may make a difference in naturalness of idiom and may also provide some incidental evidence as to what is going on in the speaker's mind, but it is incapable of making the difference between truth and falsehood of the compound. The difference in meaning between 'and', 'but', and 'although' is rhetorical, not logical. Logical notation, unconcerned with rhetorical distinctions, expresses conjunction uniformly.

For a further example of the reduction of manifold idioms of ordinary language to uniformity in logical notations, consider the idiomatic variants of 'if-then':

If p then q, p only if q, q if p, q provided that p, q in case p.

The notation '$p \rightarrow q$', insofar as it may be admitted as a version of 'if p then q' at all, is a version at once of all those variant idioms.

Note that the antecedent of a conditional, corresponding to the 'p' of '$p \rightarrow q$', is not always the part which comes first in the vernacular. It is the part rather that is governed by 'if' (or by 'in case', 'provided that', etc.), regardless of whether it comes early or late in the conditional. Thus it is that 'p if q' goes over into '$q \rightarrow p$', not '$p \rightarrow q$'. But whereas 'if' is thus ordinarily a sign of the antecedent, the attachment of 'only' reverses it; 'only if' is a sign of the consequent. Thus 'p only if q' means, not 'p if q', but 'if p then q'; not '$q \rightarrow p$', but '$p \rightarrow q$'. E.g., 'You will graduate only if your bills have been paid' does not mean 'If your bills have been paid you will graduate'; it means 'If you will graduate, your bills (will) have been paid'.

The reader may have found 'if p then q' awkward as a pronunciation of '$p \rightarrow q$', because of the separation of 'if' from 'then'. If so, the above observation on 'only if' deserves special attention; '$\rightarrow$' may be read 'only if'.

It is particularly to be noted that 'only if' does not have the sense of '$\leftrightarrow$', which is '*if and* only if'. As the words suggest, 'p if and only if q' is a conjunction of 'p if q' and 'p only if q'—hence of '$q \rightarrow p$' and '$p \rightarrow q$'.

Among the linguistic variants of 'if p then q' listed above, one more might have been included: 'not p unless q'. This variant leads to the following curious reflection: if 'not p unless q' means '$p \rightarrow q$', and '$p \rightarrow q$' means '$\bar{p} \vee q$', then 'not p unless q' must mean '$\bar{p} \vee q$', which makes 'unless' answer to '$\vee$' and hence to 'or'. Whatever strangeness there may be in equating 'unless' to 'or' is precisely the strangeness of equating 'if-then' to '$\rightarrow$'. It is sometimes felt that 'if-then' suggests a causal connection, or the like; and, insofar as it does, so also does 'unless'. But when we distill a truth function out of 'if-then' we have '$\rightarrow$', and when we distill a truth function out of 'unless' we have '$\vee$', 'or'.

The evident commutativity of 'or', i.e., the equivalence of 'p or q' with 'q or p', is less evident with 'unless'. The statements:

(1) Smith will sell unless he hears from you,
(2) Smith will hear from you unless he sells

seem divergent in meaning. However, this divergence may be attributed in part to a subtle tendency in 'unless' compounds to mention the earlier event last when time relationships are important. Because of this tendency, we are likely to construe the vague 'hears from you' in (1) as meaning 'hears from you that he should not sell', and in (2) as meaning 'hears from you that he should have sold'. But if we are to compare (1) and (2) as genuine compounds of statements, we must first render each component unambiguous and durable in its meaning—if not absolutely, at least sufficiently to exclude shifts of meaning within the space of the comparison. Thus we should perhaps revise (1) and (2) to read:

Smith will sell unless you restrain him,
Smith will be reprimanded by you unless he sells,

and so consider them to be related not as 'p unless q' and 'q unless p', but merely as 'p unless q' and 'r unless p'.

Thus far we have been surveying in a cursory way that aspect of paraphrasing which turns on mere vocabulary. We have been correlating connective words of ordinary language with the connective symbols of

symbolic logic. The last example, however, has brought to light another and subtler aspect of the task of paraphrasing: on occasion we must not only translate connectives but also rephrase the component clauses themselves, to the extent anyway of insuring them against material shifts of meaning within the space of the argument in hand. The necessity of this operation is seen more simply and directly in the following example. The two conjunctions:

(3) He went to Pawcatuck and I went along,
(4) He went to Saugatuck but I did not go along

may both be true; yet if we represent them as of the forms 'pq' and '$r\bar{q}$', as seems superficially to fit the case, we come out with an inconsistent combination '$pqr\bar{q}$'. Actually of course the 'I went along' in (3) must be distinguished from the 'I went along' whose negation appears in (4); the one is 'I went along to Pawcatuck' and the other is 'I went along to Saugatuck'. When (3) and (4) are completed in this fashion they can no longer be represented as related in the manner of 'pq' and '$r\bar{q}$', but only in the manner of 'pq' and '$r\bar{s}$'; and the apparent inconsistency disappears. In general, the trustworthiness of logical analysis and inference depends on our not giving one and the same expression different interpretations in the course of the reasoning. Violation of this principle was known traditionally as the *fallacy of equivocation*.

Insofar as the interpretation of ambiguous expressions depends on circumstances of the argument as a whole—speaker, hearer, scene, date, and underlying problem and purpose—the fallacy of equivocation is not to be feared; for, those background circumstances may be expected to influence the interpretation of an ambiguous expression uniformly wherever the expression recurs in the course of the argument. This is why words of ambiguous reference such as 'I', 'you', 'here', 'Smith', and 'Elm Street' are ordinarily allowable in logical arguments without qualification; their interpretation is indifferent to the logical soundness of an argument, provided merely that it stays the same throughout the space of the argument.

The fallacy of equivocation arises rather when the interpretation of an ambiguous expression is influenced in varying ways by immediate contexts, as in (3) and (4), so that the expression undergoes changes of meaning within the limits of the argument. In such cases we have to rephrase before proceeding; not rephrase to the extent of resolving all

ambiguity, but to the extent of resolving such part of the ambiguity as might, if left standing, end up by being resolved in dissimilar ways by different immediate contexts within the proposed logical argument. The logical connectives by which components are joined in compounds must be thought of as insulating each component from whatever influences its neighbors might have upon its meaning; each component is to be wholly on its own, except insofar as its meaning may depend on those broader circumstances which condition the meanings of words in the compound as a whole or in the logical argument as a whole.

It often becomes evident, when this warning is borne in mind, that a compound which superficially seems analyzable in terms merely of conjunction and negation really calls for logical devices of a more advanced nature. The statement:

(5) We saw Stromboli and it was erupting

is not adequately analyzed as a simple conjunction, for the construction 'was . . .-ing' in the second clause involves an essential temporal reference back to the first clause. A more adequate analysis would construe (5) rather as:

> Some moment of our seeing Stromboli was a moment of its erupting,

which involves logical structures taken up in Part II.

The general enterprise of paraphrasing statements so as to isolate their logical structures has, we have thus far seen, two aspects: the direct translating of appropriate words into logical symbols (comprising just truth-functional symbols at this level), and the rephrasing of component clauses to circumvent the fallacy of equivocation. Now a third aspect, of equal importance with the other two when our examples are of any considerable complexity, is determination of how to organize paraphrased fragments properly into a structured whole. Here we face the problem of determining the intended grouping. A few clues to grouping in statements of ordinary language have been noted (Chapter 4), but in the main we must rely on our good sense of everyday idiom for a sympathetic understanding of the statement and then re-think the whole in logical symbols. When a statement is complex, it is a good plan to look for the outermost structure first and then *paraphrase inward,* step by step. This procedure

has the double advantage of dividing the problem up into manageable parts, and of keeping the complexities of grouping under control. E.g., consider the statement:

(6) If Jones is ill or Smith is away then neither will the Argus deal be concluded nor will the directors meet and declare a dividend unless Robinson comes to his senses and takes matters into his own hands.

First we seek the main connective of (6). Reasoning as in Chapter 4, we can narrow the choice down to 'if-then' and 'unless'; suppose we decide on 'if-then'. The outward structure of (6), then, is that of a conditional; so let us impose just this much structure explicitly upon (6), postponing minuter analysis. We have:

(7) Jones is ill or Smith is away $\rightarrow$ neither will the Argus deal be concluded nor will the directors meet and declare a dividend unless Robinson comes to his senses and takes matters into his own hands.

Next we may consider, as if it were a separate problem removed from (7), just the long compound 'neither . . . hands'. We decide, let us suppose, that its main connective is 'unless'. Treating 'unless' as '$\vee$', we turn (7) as a whole into:

(8) Jones is ill or Smith is away $\rightarrow .$ neither will the Argus deal be concluded nor will the directors meet and declare a dividend $\vee$ Robinson will come to his senses and take matters into his own hands.

Now we take up, as if it were a separate problem removed from (8), the longest component not yet analyzed; viz., 'neither . . . dividend'. The main connective here is clearly 'neither-nor'. Reflecting then that 'neither r nor s' in general goes into symbols as '$\bar{r}\bar{s}$', we rewrite 'neither . . . dividend' accordingly; (8) thus becomes:

(9) Jones is ill or Smith is away $\rightarrow .$ $-$(the Argus deal will be concluded) $-$(the directors will meet and declare a divi-

dend) ∨ Robinson will come to his senses and take matters into his own hands.

Directing our attention finally to the various short compounds which remain unanalyzed in (9), we turn the whole into:

(10) Jones is ill ∨ Smith is away .→: −(the Argus deal will be concluded) −(the directors will meet . the directors will declare a dividend) ∨. Robinson will come to his senses . Robinson will take matters into his own hands.

Put schematically, the total structure is:

(11) $$p \vee q \;.\!\rightarrow .\; \bar{r} - (st) \vee uv.$$

EXERCISES

1. Justify inference of the conclusion:

   If Smith is away and Robinson does not come to his senses then the Argus deal will not be concluded

   from (6). Method: Find the schema which corresponds to this conclusion as (11) does to (6); then show that this schema is implied by (11).

2. Determine which of these statements implies which:

   Jones is not eligible unless he has resigned his commission and signed a waiver.
   Jones is eligible if he has resigned his commission or signed a waiver.
   Jones is eligible only if he has signed a waiver.

   Method: Paraphrase the statements, represent their structure schematically, and test the schemata. Show all steps.

3. Paraphrase inward, showing and justifying each step:

   If either the Giants or the Bruins win and the Jackals take

second place, then I'll recover past losses and either buy a clavichord or fly to Barbuda.

4. Paraphrase inward, showing and justifying each step:

If the tree rings have been correctly identified and the mace is indigenous, then the Ajo culture antedated the Tula if and only if the Tula culture was contemporary with or derivative from that of the present excavation.

9
EQUIVALENCE

Two truth-functional schemata are called *equivalent* if they agree with each other in point of truth value under every interpretation of their letters, or in other words if they agree case by case under truth-value analysis. In anticipation, various cases of equivalence were noted in Chapters 1–3:

'p' to '$\bar{\bar{p}}$', 'pp', and '$p \vee p$',
'pq' to 'qp' and '$-(\bar{p} \vee \bar{q})$',
'$p \vee q$' to '$q \vee p$' and '$-(\bar{p}\bar{q})$',
'$p \rightarrow q$' to '$-(p\bar{q})$' and '$\bar{p} \vee q$',
'$p \leftrightarrow q$' to '$p \rightarrow q \,.\, q \rightarrow p$' and '$-(p\bar{q})\ -(q\bar{p})$'.

'$pq \,.\, r$' to '$p \,.\, qr$',
'$p \vee q \,.\vee\, r$' to '$p \vee .\, q \vee r$',
'$-(pq)$' to '$\bar{p} \vee \bar{q}$',
'$-(p \vee q)$' to '$\bar{p}\bar{q}$',

To test two schemata for equivalence, we might make truth-value analyses of the two schemata and see if they agree case by case. But there is another way which tends to be easier: we may form a biconditional of the two schemata, and test it for validity. For, according to our definition, two schemata S_1 and S_2 are equivalent if and only if no interpretation makes S_1 and S_2 unlike in truth value; hence if and only if no interpretation falsifies the biconditional whose sides are S_1 and S_2. Thus, just as implication is validity of the conditional, so equivalence is validity of the biconditional.

To determine the equivalence of '$p \,.\, q \vee r$' to '$pq \vee pr$', e.g., we check the validity of the corresponding biconditional:[12]

$$p \,.\, q \vee r \;.\leftrightarrow.\; pq \vee pr$$

$$\top \,.\, q \vee r \;.\leftrightarrow.\; \top q \vee \top r \qquad\qquad \bot \,.\, q \vee r \;.\leftrightarrow.\; \bot q \vee \bot r$$

$$q \vee r \;.\leftrightarrow.\; q \vee r \qquad\qquad \top$$

$$\top$$

In similar fashion it may be checked that 'p' is equivalent to each of:

(1) $\bar{\bar{p}}$, pp, $p \vee p$, $p \vee pq$, $p \,.\, p \vee q$, $pq \vee p\bar{q}$, $p \vee q \,.\, p \vee \bar{q}$.

It has been said that the most conspicuous purpose of logic, in its application to ordinary discourse, is the justification and criticism of inference. But a second purpose, almost as important, is transformation of statements. It is often desirable to transform one statement into another which "says the same thing" in a different form—a form which is simpler, perhaps, or more convenient for the particular purposes in hand. Now insofar as such transformations are justifiable by considerations purely of truth-functional structure (rather than turning upon other sorts of logical structure which lie beyond the scope of Part I), a technique for their justification is at hand in our test of equivalence of truth-functional schemata. Transformation, e.g., of:

(2) The admiral will speak and either the dean or the president will introduce him

into:

(3) Either the admiral will speak and the dean will introduce him or the admiral will speak and the president will introduce him,

or vice versa, is justified by the equivalence of '$p \,.\, q \vee r$' to '$pq \vee pr$', and this equivalence is verified by mechanical test as above. The state-

[12] In pursuance of the policy announced in Chapter 5, all intermediate steps of resolution in this analysis are left to the reader to fill in. The reason an intermediate stage '$q \vee r \;.\leftrightarrow.\; q \vee r$' is shown in the left-hand part of the analysis is that the passage from this to '$\top$' is not by resolution but by the rule of patently valid clauses, Chapter 6.

ments (2) and (3) may, by an extension of terminology similar to that made in Chapter 7, be spoken of as truth-functionally equivalent.

It is evident from our definitions and testing techniques that

(i) Equivalence is mutual implication.

From this law and (i)–(iv) of Chapter 7, these clearly follow:

(ii) Any schema is equivalent to itself.

(iii) If one schema is equivalent to a second and the second is equivalent to a third then the first is equivalent to the third.

(iv) If one schema is equivalent to a second then the second is equivalent to the first. (Not so for implication!)

(v) Valid schemata are equivalent to one another and to no others; and similarly for inconsistent schemata.

Substitution was observed in Chapter 6 to preserve validity. Since implication and equivalence are merely validity of a conditional and a biconditional, it follows that substitution also preserves implication and equivalence. From the equivalence of 'p' to each of the schemata in (1), e.g., we may infer by substitution that '$\bar{r}$' is equivalent to each of '$\bar{\bar{r}}$', '$\bar{r}\bar{r}$', '$\bar{r} \vee \bar{r}$', '$\bar{r} \vee \bar{r}s$', etc.; also that '$\bar{q}r$' is equivalent to each of '$- -(\bar{q}r)$', '$\bar{q}r\bar{q}r$', '$\bar{q}r \vee \bar{q}r$', etc.; and correspondingly for any other substitution upon 'p' and 'q' in (1). The particular family of equivalences thus generated will be used later as a means of simplifying schemata.

Appeal to substitution is helpful incidentally in justifying these two convenient ways of describing implication in terms of equivalence:

(vi) S_1 implies S_2 if and only if S_1 is equivalent to the conjunction of S_1 and S_2.

(vii) S_1 implies S_2 if and only if S_2 is equivalent to the alternation of S_1 and S_2.

To justify (vi), observe to begin with that '$p \rightarrow q$' and '$p \leftrightarrow pq$' are equivalent by truth-value analysis. It follows by substitution, then, that the conditional formed of S_1 and S_2 is equivalent to the biconditional formed of S_1 and the conjunction of S_1 and S_2. So, by (v), that conditional is valid if and only if this biconditional is valid. But validity of the conditional is implication, and validity of the biconditional is equiva-

lence; so (vi) follows. The justification of (vii) proceeds similarly from the verifiable equivalence of '$p \rightarrow q$' and '$q \leftrightarrow .\ p \vee q$'.

The following two laws have to do only with implication, but I have deferred them to here because they are obtained easily from (vi) and (vii).

(viii) S implies each of S_1 and S_2 if and only if S implies the conjunction of S_1 and S_2.

(ix) S_1 and S_2 each imply S if and only if the alternation of S_1 and S_2 implies S.

Justification of (viii): Suppose S implies S_1 and S_2. Then S is equivalent, by (vi), to the conjunction of S and S_1, and also to the conjunction of S and S_2. So S is equivalent to the conjunction of all three. But, by (vi), this is the same as saying that S implies the conjunction of S_1 and S_2. So we have seen that, if S implies S_1 and also S_2, it implies their conjunction. The converse is trivial.

A parallel argument for (ix), based on (vii), is left to the reader.

The laws (viii) and (ix) afford a convenient extension of the method of fell swoop. E.g., to see whether '$pq \vee \bar{p}r \vee q\bar{r}$' implies a given schema, we have only (thanks to (ix)) to see whether each of 'pq', '$\bar{p}r$', and '$q\bar{r}$' individually implies it; and this may be seen by three fell swoops. Again, to see whether a given schema implies '$p \vee q\ .\ p \rightarrow r\ .\ q \vee \bar{r}$', we have only (thanks to (viii)) to see whether it implies each of '$p \vee q$', '$p \rightarrow r$', and '$q \vee \bar{r}$'; and this again may be seen by three fell swoops.

We must not hope to build (viii) and (ix) out into a square by supposing that S implies one or the other of S_1 and S_2 whenever S implies their alternation; nor, again, that S_1 or S_2 implies S whenever their conjunction implies S. After all, 'p' implies '$pq \vee p\bar{q}$' without implying either 'pq' or '$p\bar{q}$'; and 'p' is implied by '$p \vee q\ .\ p \vee \bar{q}$' but neither by '$p \vee q$' nor by '$p \vee \bar{q}$'.

Substitution consists always in putting schemata for single letters, and for all recurrences of the letters. When these restrictions are not met, the putting of one schema for another will be called not substitution but *interchange*. Thus interchange consists in putting one schema for another which need not be a single letter, and which need not be supplanted in all its recurrences. What has been said of substitution, that it preserves implication, equivalence, and inconsistency, cannot of course be said in general of interchange. But there are useful laws of interchange, the least of which is this *first law of interchange:* Think of '$\ldots p \ldots$' as any

schema containing 'p', and of '$\ldots q \ldots$' as formed from '$\ldots p \ldots$' by putting 'q' for one or more occurrences of 'p'; then

$$\text{'}p \leftrightarrow q\text{'} \quad \text{implies} \quad \text{'}\ldots p \ldots . \leftrightarrow . \ldots q \ldots\text{'}.$$

(Similarly for any other letters instead of 'p' and 'q'.) Let us see why the law holds. We want to show that any interpretation of letters which makes '$p \leftrightarrow q$' come out true will make '$\ldots p \ldots . \leftrightarrow . \ldots q \ldots$' come out true. But to make '$p \leftrightarrow q$' come out true we must either put '$\top$' for both 'p' and 'q' or else '$\bot$' for both 'p' and 'q'; and in either case '$\ldots p \ldots$' and '$\ldots q \ldots$', which differed only in 'p' and 'q', become indistinguishable from each other, so that their biconditional reduces to '$\top$'.

Now we can establish a more important *second law of interchange:* If S_1 and S_2 are equivalent, and S_2' is formed from S_1' by putting S_2 for one or more occurrences of S_1, then S_1' and S_2' are equivalent. E.g., this law enables us to argue from the equivalence of '$p \rightarrow q$' and '$-(p\bar{q})$' to the equivalence of '$p \rightarrow q \,.\vee\, r$' and '$-(p\bar{q}) \vee r$'. The rough idea is, in school jargon, that putting equals for equals yields equals.

This second law of interchange is established as follows. Choose any two letters not appearing in S_1' nor in S_2'. They are, let us imagine, 'p' and 'q'. Then put 'p' for the occurrences of S_1 in question in S_1'; the result may be represented as '$\ldots p \ldots$', and the result of similarly using 'q' may be represented as '$\ldots q \ldots$'. By the first law of interchange, '$p \leftrightarrow q$' implies '$\ldots p \ldots . \leftrightarrow . \ldots q \ldots$'. By substitution of S_1 for 'p' and S_2 for 'q' in this implication, we may conclude that the biconditional of S_1 and S_2 implies the biconditional of S_1' and S_2'. But the biconditional of S_1 and S_2 is valid, since S_1 and S_2 are equivalent. Therefore the biconditional of S_1' and S_2' is valid; see (iv) of Chapter 7. Therefore S_1' and S_2' are equivalent.

This second law assures us that we can interchange equivalents S_1 and S_2 in any schema S_1' without affecting the outcome of a truth-value analysis; for, S_1' and the result S_2' will be equivalent, and equivalent schemata are schemata that agree case by case under truth-value analysis. There thus follows this *third law of interchange:* Interchange of equivalents preserves validity, implication, equivalence, and inconsistency; and, unlike substitution for letters, it even preserves consistency, nonvalidity, nonimplication, and nonequivalence.

Substitution for letters must, we saw, be construed as uniform and exhaustive; but there is no such requirement in the case of interchanging equivalents. If in the valid schema '$p \vee \bar{p}$' we substitute 'qr' for 'p', we may infer the validity of '$qr \vee -(qr)$' and this only; but if in that same

valid schema '$p \vee \bar{p}$' we elect rather to put 'pp' for its *equivalent* 'p', we are entitled thereby to infer the validity not merely of '$pp \vee -(pp)$', but equally of '$pp \vee \bar{p}$' and '$p \vee -(pp)$'.

Since interchange of equivalents does not affect the outcome of a truth-value analysis, it proves to be a convenient adjunct to the technique of truth-value analysis; for, if we supplant schemata by simpler equivalents in the course of such analyses, our computations are reduced. In particular, accordingly, whenever a configuration of any of the seven forms depicted in (1) makes its appearance in the course of a truth-value analysis, let us immediately simplify it before proceeding. We are not only to put 'p' for its equivalents '$\bar{\bar{p}}$', 'pp', '$p \vee p$', '$p \vee pq$', etc., but correspondingly '$\bar{r}$' for '$\bar{\bar{\bar{r}}}$', '$\bar{r}\bar{r}$', '$\bar{r} \vee \bar{r}$', '$\bar{r} \vee \bar{r}s$', etc., and '$\bar{q}r$' for '$- -(\bar{q}r)$', '$\bar{q}r\bar{q}r$', '$\bar{q}r \vee \bar{q}r$', etc.

With our new policy in mind let us take another turn at the long schema which was analyzed in Chapter 6:

$$p \vee q \,.\, p \vee \bar{q} \,.\vee\, \bar{p}q :\leftrightarrow q \,.:\!\rightarrow.\, pr \vee p\bar{r}$$

$$p \vee \bar{p}q \,.\leftrightarrow q :\!\rightarrow p$$

$$\top \vee \bot q \,.\leftrightarrow q :\!\rightarrow \top \qquad \bot \vee \top q \,.\leftrightarrow q :\!\rightarrow \bot$$

$$\top \qquad\qquad -(q \leftrightarrow q)$$

$$\bot$$

Here the original schema is subjected to some simplifications before the substitution of signs for 'p' is even begun. The simplifications consist in reducing both '$p \vee q \,.\, p \vee \bar{q}$' and '$pr \vee p\bar{r}$' to '$p$'; for '$p \vee q \,.\, p \vee \bar{q}$' is the last of the schemata in (1), and '$pr \vee p\bar{r}$' is the next to last with 'r' substituted for 'q'.

Next let us turn back to the first long truth-value analysis of Chapter 6. Under the new procedure it would run rather thus:

$$pq \vee p\bar{r} \vee \bar{p}r \vee \bar{p}s \vee \bar{q}r \vee \bar{r}\bar{s}$$

$$\top q \vee \top \bar{r} \vee \bot r \vee \bot s \vee \bar{q}r \vee \bar{r}\bar{s} \qquad \bot q \vee \bot \bar{r} \vee \top r \vee \top s \vee \bar{q}r \vee \bar{r}\bar{s}$$

$$q \vee \bar{r} \vee \bar{q}r \vee \bar{r}\bar{s} \qquad r \vee s \vee \bar{q}r \vee \bar{r}\bar{s}$$

$$q \vee \bar{r} \vee \bar{q}r \qquad r \vee s \vee \bar{r}\bar{s}$$

$$\top \vee \bar{r} \vee \bot r \qquad \bot \vee \bar{r} \vee \top r \qquad \top \vee s \vee \bot\bar{s} \qquad \bot \vee s \vee \top\bar{s}$$

$$\top \qquad \bar{r} \vee r \qquad \top \qquad s \vee \bar{s}$$

$$\top \qquad\qquad \top$$

In this case none of the seven forms listed in (1) is visible in the original schema as it stands, but some emerge as the analysis proceeds. On the left side, '$q \vee \bar{r} \vee \bar{q}r \vee \bar{r}\bar{s}$' is reduced to '$q \vee \bar{r} \vee \bar{q}r$' by putting '$\bar{r}$' for

'$\bar{r} \vee \bar{r}\bar{s}$'. On the right side, similarly, '$r \vee s \vee \bar{q}r \vee \bar{r}\bar{s}$' is reduced to '$r \vee s \vee \bar{r}\bar{s}$' by putting '$r$' for '$r \vee \bar{q}r$'.

Both of the simplifications last noted are based on the equivalence of '$p \vee pq$' to 'p'; but they involve also a mental switching of conjunctions and alternations. The clause '$\bar{r} \vee \bar{r}\bar{s}$' which is to give way to '$\bar{r}$' is not even visible in '$q \vee \bar{r} \vee \bar{q}r \vee \bar{r}\bar{s}$' until we think of the part '$\bar{q}r \vee \bar{r}\bar{s}$' as switched to read '$\bar{r}\bar{s} \vee \bar{q}r$'; nor is the clause '$r \vee \bar{q}r$' visible in '$r \vee s \vee \bar{q}r \vee \bar{r}\bar{s}$' until we think of the part '$s \vee \bar{q}r$' as switched to read '$\bar{q}r \vee s$'. Even when this clause '$r \vee \bar{q}r$' has been isolated, moreover, its equivalence to 'r' is not inferred from the equivalence of '$p \vee pq$' to 'p' *merely* by substitution; we have also, mentally, to reread '$r \vee \bar{q}r$' as '$r \vee r\bar{q}$' by switching the conjunction. Such preparatory switching of alternations and conjunctions involves a tacit appeal to further equivalences: the equivalence of '$p \vee q$' to '$q \vee p$' and of 'pq' to 'qp'. But these steps drop out of consciousness if we school ourselves, as we well may, to disregard typographical order among the components of a conjunction and of an alternation.

It is arbitrary to single out just these seven equivalences, viz., the equivalence of 'p' to each of the seven schemata in (1), as a basis for simplifications auxiliary to truth-value analyses. A further convenient equivalence, which could in fact have been exploited in both of the truth-value analyses last set forth, is the equivalence of '$p \vee \bar{p}q$' to '$p \vee q$'. Another convenient one is the equivalence of '$p \,.\, \bar{p} \vee q$' to 'pq'. The practical investigator will use any simplificatory equivalences that occur to him. For the standardizing of exercises, a convenient compromise might be to allow use of the seven equivalences singled out in (1) and the further ones assembled at the beginning of the present chapter.

Pure biconditionals, that is, schemata built up of statement letters by means solely of '$\leftrightarrow$', admit of a startlingly simple test of equivalence. They are equivalent if and only if just the same letters occur an odd number of times in each. A pure biconditional is valid if and only if no letter occurs in it an odd number of times. Negation signs, moreover, can be sprinkled through a pure biconditional without any effect if they are even in number. These observations are due to Leśniewski (pp. 26, 29).

HISTORICAL NOTE: The confusion between implication and the conditional, deplored at the end of Chapter 7, has carried with it a confu-

sion between equivalence and the biconditional. The deplorable habit still persists of pronouncing '↔' as 'is equivalent to' instead of as 'if and only if'.

EXERCISES

1. Determine which of these are equivalent to '$pq \to r$' and which to '$p \vee q \mathrel{.\!\to} r$':

   $$p \to .\, q \to r, \quad q \to .\, p \to r, \quad p \to r \,.\, q \to r, \quad p \to r \mathrel{.\vee.} q \to r.$$

2. Determine which of these is equivalent to '$p \to qr$' and which to '$p \to .\, q \vee r$':

   $$p \to q \,.\, p \to r, \qquad p \to q \mathrel{.\vee.} p \to r.$$

3. Determine any equivalent pairs from among these:

   $$p \to q, \quad \bar{p} \to \bar{q}, \quad \bar{p} \to q, \quad q \to p, \quad \bar{q} \to \bar{p}, \quad \bar{q} \to p.$$

   This means fifteen short tests.

4. Making full use of the new simplification procedure, test each of the following three pairs for equivalence by truth-value analysis of biconditionals:

   $pq \vee pr \vee qr$, $\quad p \vee q \,.\, p \vee r \,.\, q \vee r$;

   $pqr \vee pqs \vee prs \vee qrs$, $\quad p \vee q \vee r \,.\, p \vee q \vee s \,.\, p \vee r \vee s \,.\, q \vee r \vee s$;

   $pqr \vee pq\bar{r} \vee p\bar{q}r \vee p\bar{q}\bar{r}$, $\quad p \vee q \vee r \,.\, p \vee q \vee \bar{r} \,.\, p \vee \bar{q} \vee r \,.\, p \vee \bar{q} \vee \bar{r}$.

5. By the same method, check the schemata:

   $$p \to .\, q \leftrightarrow r, \qquad pq \leftrightarrow pr, \qquad p \vee q \mathrel{.\!\leftrightarrow.} p \vee r$$

   for equivalence each with each.

6. Check these for equivalence by the same method:

   $$p \to q \,.\, q \to r \,.\, r \to p, \qquad p \leftrightarrow r \,.\, q \leftrightarrow r.$$

7. See if the biconditional is associative.

8. Justify (ix).

10
ALTERNATIONAL NORMAL SCHEMATA

The notations '∨', '→', and '↔' are superfluous, we know, in that all use of them can be paraphrased into terms of conjunction and negation. The sign '→', however, has been seen to have a special utility in the testing of implication; for, to test implication we form a conditional (with help of '→') and test its validity. The sign '↔' has been seen to be of similar use in the testing of equivalence. So there is good reason for having added the strictly superfluous signs '→' and '↔'. Now the advantages of retaining '∨' are of quite a different kind, and will become evident in the course of the present chapter and the next two.

What are known as DeMorgan's laws affirm the equivalence of

(i) '$-(p \vee q \vee \ldots \vee s)$' to '$\bar{p}\bar{q}\ldots\bar{s}$'

and

(ii) '$-(pq\ldots s)$' to '$\bar{p} \vee \bar{q} \vee \ldots \vee \bar{s}$'.

For the case of just 'p' and 'q', these laws were already noted in Chapter 1. The further cases follow by substitution and interchange. E.g., from the equivalence of '$-(p \vee q)$' to '$\bar{p}\bar{q}$' we have, by substitution of '$p \vee q$' for 'p' and 'r' for 'q', the equivalence of '$-(p \vee q \vee r)$' to '$-(p \vee q)\bar{r}$'; and thence, putting '$\bar{p}\bar{q}$' for its equivalent '$-(p \vee q)$', we obtain the equivalence of '$-(p \vee q \vee r)$' to '$\bar{p}\bar{q}\bar{r}$'.

DeMorgan's laws are useful in enabling us to avoid negating conjunctions and alternations. We never need apply negation to the whole of an alternation, since '$-(p \vee q \vee \ldots \vee s)$' is equivalent to '$\bar{p}\bar{q}\ldots\bar{s}$'; and we never need apply negation to the whole of a conjunction, since '$-(pq\ldots s)$' is equivalent to '$\bar{p} \vee \bar{q} \vee \ldots \vee \bar{s}$'. Also of course we never need apply negation to a negation, since '$\bar{\bar{p}}$' is equivalent to 'p'. For that matter, we also never need apply negation to a conditional or biconditional; for, by the method of the preceding chapter it is easy to verify the equivalence of

(iii) $\quad$ '$-(p \to q)$' $\quad$ to $\quad$ '$p\bar{q}$'

and of

(iv) $\quad$ '$-(p \leftrightarrow q)$' $\quad$ to $\quad$ '$\bar{p} \leftrightarrow q$' $\quad$ and to $\quad$ '$p \leftrightarrow \bar{q}$'.

So any truth-functional schema can be put over into an equivalent in which negation never applies to anything but individual letters. Transformation of this kind is generally conducive to easy intelligibility.

E.g., consider the forbidding schema:

(1) $\quad -\{p \to \bar{s}q \mathrel{.\!\to} -(sq \to p) : -[-(rp) - (p \to \bar{s})]\}.$

(It is of some help to vary parentheses thus with brackets and braces when they are deeply nested.) Now since (1) has the outward form '$-(tu)$', it can be transformed by (ii) to read:

$$-[p \to \bar{s}q \mathrel{.\!\to} -(sq \to p)] \vee --[-(rp) - (p \to \bar{s})].$$

Cancellation of '$--$' reduces this to:

(2) $\quad -[p \to \bar{s}q \mathrel{.\!\to} -(sq \to p)] \vee -(rp) - (p \to \bar{s}).$

Then by (iii) we transform the first half of (2) into:

$$p \to \bar{s}q \,.\, --(sq \to p),$$

or

$$p \to \bar{s}q \,.\, sq \to p,$$

so that (2) becomes:

(3) $\quad p \to \bar{s}q \,.\, sq \to p \mathrel{.\vee} -(rp) - (p \to \bar{s}).$

By (ii) again, '$-(rp)$' here becomes '$\bar{r} \vee \bar{p}$', and, by (iii) again, '$-(p \to \bar{s})$' becomes '$p\bar{\bar{s}}$' or 'ps', so that (3) comes down to:

(4) $\quad p \to \bar{s}q \,.\, sq \to p \mathrel{.\vee.} \bar{r} \vee \bar{p} \,.\, ps,$

in which, finally, all negation signs are limited to single letters. (4) is far easier to grasp than (1).

Such is the advantage of confining negation to single letters. Now it will be found in general that still further perspicuity can be gained by confining conjunction to single letters and negations of letters; and it will be found also that such confinement of conjunction can, like the confinement of negation to single letters, always be accomplished. The law which makes this possible is known as the *law of distributivity of conjunction through alternation,* and runs as follows:

'$p \,.\, q \vee r \vee \ldots \vee t$' is equivalent to '$pq \vee pr \vee \ldots \vee pt$'.

Regardless of the number of letters involved, the equivalence is readily verified by the method of the preceding chapter:

$$p \,.\, q \vee r \vee \ldots \vee t \;.\leftrightarrow.\; pq \vee pr \vee \ldots \vee pt$$

$$\top \,.\, q \vee r \vee \ldots \vee t \;.\leftrightarrow.\; \top q \vee \top r \vee \ldots \vee \top t \qquad \bot \,.\, q \vee r \vee \ldots \vee t \;.\leftrightarrow.\; \bot q \vee \bot r \vee \ldots \vee \bot t$$

$$q \vee r \vee \ldots \vee t \;.\leftrightarrow.\; q \vee r \vee \ldots \vee t \qquad\qquad \bot \leftrightarrow \bot$$

$$\top \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad \top$$

This law, like the familiar identity:

$$x(y + z + \ldots + w) = xy + xz + \ldots + xw$$

of algebra, authorizes the convenient operation of "multiplying out." Thanks to it, we need never acquiesce in a conjunction which has an alternation as component; we can always distribute the other part of the conjunction through the alternation, as above, so as to come out with an alternation of simpler conjunctions.

Since order is immaterial to conjunction, such distribution can be worked equally well in reverse: not only is '$p \,.\, q \vee r \vee \ldots \vee t$' equivalent to '$pq \vee pr \vee \ldots \vee pt$', but also '$q \vee r \vee \ldots \vee t \,.\, p$' is equivalent to '$qp \vee rp \vee \ldots \vee tp$'. These two sorts of distribution are indeed one and the same, once we learn to ignore order of conjunction.

When we have a conjunction of two alternations, distribution takes the form of the familiar "cross-multiplying" of algebra; e.g., '$p \vee t \,.\, q \vee r \vee s$' comes out '$pq \vee pr \vee ps \vee tq \vee tr \vee ts$'. For, we begin by handling '$q \vee r \vee s$' as we might a single letter 'u'; thus, just as '$p \vee t \,.\, u$' would become '$pu \vee tu$' by (reverse) distribution, so '$p \vee t \,.\, q \vee r \vee s$' becomes '$p \,.\, q \vee r \vee s \;.\vee.\; t \,.\, q \vee r \vee s$'. Afterward,

distribution of 'p' turns the part '$p \cdot q \vee r \vee s$' into '$pq \vee pr \vee ps$', and distribution of 't' turns the part '$t \cdot q \vee r \vee s$' into '$tq \vee tr \vee ts$'.

Let us now go back to (4) and improve it by distributing. We thereby change the part '$\bar{r} \vee \bar{p} \cdot ps$' of (4) to '$\bar{r}ps \vee \bar{p}ps$', so that (4) becomes:

(5) $$p \rightarrow \bar{s}q \cdot sq \rightarrow p \;.\vee\; \bar{r}ps \vee \bar{p}ps.$$

We can open the way to further distribution if we get rid of '$\rightarrow$', translating '$t \rightarrow u$' in general as '$\bar{t} \vee u$'. Such translation turns (5) into:

$$\bar{p} \vee \bar{s}q \cdot -(sq) \vee p \;.\vee\; \bar{r}ps \vee \bar{p}ps,$$

which, when '$-(sq)$' is changed to '$\bar{s} \vee \bar{q}$' by (ii), becomes:

$$\bar{p} \vee \bar{s}q \cdot \bar{s} \vee \bar{q} \vee p \;.\vee\; \bar{r}ps \vee \bar{p}ps.$$

Now the part '$\bar{p} \vee \bar{s}q \cdot \bar{s} \vee \bar{q} \vee p$' can be ''cross-multiplied,'' so that the whole becomes:

(6) $$\bar{p}\bar{s} \vee \bar{p}\bar{q} \vee \bar{p}p \vee \bar{s}q\bar{s} \vee \bar{s}q\bar{q} \vee \bar{s}qp \vee \bar{r}ps \vee \bar{p}ps.$$

We can quickly shorten this result by deleting the patently inconsistent clauses '$\bar{p}p$', '$\bar{s}q\bar{q}$', and '$\bar{p}ps$'. We then have:

(7) $$\bar{p}\bar{s} \vee \bar{p}\bar{q} \vee \bar{s}q\bar{s} \vee \bar{s}qp \vee \bar{r}ps.$$

Such deletion is a case of the procedure explained in Chapter 6: each of the patently inconsistent clauses may be thought of as supplanted by '$\perp$', which afterward drops by resolution ((ii) of Chapter 5).

Also we may drop any duplications from conjunctions—thus reducing '$\bar{s}q\bar{s}$' to '$\bar{s}q$'. This was the second of the seven forms of simplification noted in (1) of the preceding chapter. So (7) becomes:

(8) $$\bar{p}\bar{s} \vee \bar{p}\bar{q} \vee \bar{s}q \vee \bar{s}qp \vee \bar{r}ps,$$

which wears its meaning on its sleeve. This its equivalents (1) and (4) could scarcely have been said to do.

The forms (6)–(8) share the following three noteworthy properties: '$\rightarrow$' and '$\leftrightarrow$' do not occur; negation is confined to single letters; and conjunction is confined to letters and negations of letters. Schemata hav-

ing these three properties will be called *alternational normal schemata*.

This essentially negative characterization may be reformulated in more positive terms as follows. Let us speak of single letters and negations of single letters collectively as *literals;* thus 'p', 'q', '$\bar{p}$', etc. are literals. The alternational normal schemata, then, are the literals, the conjunctions of literals, the alternations of literals, the alternations of conjunctions of literals, and the alternations of literals with conjunctions of literals. This characterization can be put much more compactly if we allow ourselves to speak of conjunctions and alternations not only of two or more components but also of one component—meaning thereby the component itself. Under this usage, '$\bar{p}$', '$\bar{p}q$', and '$\bar{p}q\bar{r}$' are conjunctions respectively of one, two, and three literals; and correspondingly for alternation. An alternational normal schemata, then, is describable simply as any *alternation of conjunctions of literals*—that is, any alternation of *one* or more conjunctions of *one* or more literals. These conjunctions, whereof the alternational normal schema is an alternation, are called its *clauses*.

The process whereby (1) was transformed into its alternational normal equivalent (6) can be reproduced for all schemata. Given any schema, we can rid it of '$\rightarrow$' and '$\leftrightarrow$' by familiar translations: '$p \rightarrow q$' becomes '$\bar{p} \vee q$' and '$p \leftrightarrow q$' becomes '$pq \vee \bar{p}\bar{q}$'. We can confine negation to single letters by (i)–(iv), or simply by (i)–(ii) having first got rid of '$\rightarrow$' and '$\leftrightarrow$'. Finally we can confine conjunction to literals by persistent distribution. The reader will find by experiment that labor is generally saved, in these transformations, by working from the outside inward.

This is all there is, strictly speaking, to transformation into alternational normal form. Simplification, however, as in passing from (6) to (8), is always welcome too. This gets rid, we saw, of duplications of a literal within a clause, and it gets rid of any inconsistent clause—unless of course the whole schema has boiled down to a single inconsistent clause, say '$\bar{q}sq$', whose omission would leave us with nothing at all. Even in this extreme case a small simplification can still be made: we can write '$p\bar{p}$', since all inconsistent schemata are equivalent.

Thus in the alternational normal form we have an immediate test of inconsistency: just simplify by dropping inconsistent clauses, in the above fashion, and see if nothing but a visibly inconsistent clause remains.

Alternational normal schemata are generally convenient because their net import is so readily grasped: we can tell at a glance what interpretations will make them true. E.g., an interpretation will make (8) true if and only if it either interprets 'p' and 's' as false (making the first clause of (8)

true), or interprets 'p' and 'q' as false (making the second clause true), or interprets 's' as false and 'q' as true, or etc.

One of the simplification laws that are collected in (1) of Chapter 9 reduced '$pq \vee p\bar{q}$' to 'p'. If perversely we apply this law backwards, as a complication law, we can expand one alternational normal schema into another which has certain bold traits worth noticing. This expansive transformation is called *development;* 'p' goes into '$pq \vee p\bar{q}$' by development with respect to 'q'. Now if we persistently develop each clause of an alternational normal schema with respect to each letter that is absent from the clause, we arrive at the *developed alternational normal form.* For instance, '$\bar{s}q \vee \bar{r}ps$' becomes first:

$$\bar{s}qp \vee \bar{s}q\bar{p} \vee \bar{r}psq \vee \bar{r}ps\bar{q}$$

and eventually:

$$\bar{s}qpr \vee \bar{s}qp\bar{r} \vee \bar{s}q\bar{p}r \vee \bar{s}q\bar{p}\bar{r} \vee \bar{r}psq \vee \bar{r}ps\bar{q},$$

or, alphabetized:

$$pqr\bar{s} \vee pq\bar{r}\bar{s} \vee \bar{p}qr\bar{s} \vee \bar{p}q\bar{r}\bar{s} \vee pq\bar{r}s \vee p\bar{q}\bar{r}s.$$

This form, when thus alphabetized and freed of repetitions, is in effect a truth table; each of its clauses depicts one of the ways of assigning truth values to letters that makes the schema come out true. Implication becomes recognizable on sight: if the letters of two developed alternational normal schemata are alike, then the one schema implies the other just in case all its clauses are among the clauses of the other. The mark of validity, in developed alternational normal form, is that all possible clauses are present: all 2^n of them, where n is the number of different letters. The mark of inconsistency is disappearance; no clauses.

HISTORICAL NOTE: DeMorgan's laws were named for Augustus DeMorgan, who flourished in 1846–64; but they were known to William of Ockham five centuries earlier. (Cf. Łukasiewicz, "Zur Geschichte.") The idea of *development* and its name go back to DeMorgan's contemporary George Boole. The alternational normal form was familiar to Ernst

Schröder by 1877 and is doubtless older. This form is often called, less suggestively, the disjunctive normal form.

EXERCISES

1. By successive transformations, transform each of these schemata into an alternational normal schema.

$$-(p \vee -\{q \vee -[r \vee -(q \vee p)]\}),$$
$$p \to q \,.\, q \to r \,.\!\to.\, p \to r,$$
$$p \to q \,.\!\to p :\!\leftrightarrow \bar{p},$$
$$p \leftrightarrow q \,.\, q \leftrightarrow r,$$
$$p \leftrightarrow q \,.\!\leftrightarrow r,$$
$$pq \leftrightarrow r.$$

2. Develop each of the results of the preceding exercise into developed alternational normal form.

3. Judge whether this is a sound general method of testing alternational normal truth-functional schemata for equivalence: *decide by fell swoop whether each clause of each schema implies the other schema*. Justify your judgment.

4. Check the implications in Exercises 1–4 of Chapter 7 by the method of developed alternational normal forms.

11
SIMPLIFICATION

The perspicuity of the alternational normal form can be enhanced by pressing simplification. In passing from (7) to (8) in the preceding chapter we used one of the seven forms of simplification which were noted in connection with (1) of Chapter 9, viz. 'pp' to 'p'. But others of the seven may likewise be used to advantage, e.g., that of '$p \vee pq$' to 'p'. Thus, reducing '$\bar{s}q \vee \bar{s}qp$' to '$\bar{s}q$' in (8), we get:

(1) $\bar{p}\bar{s} \vee \bar{p}\bar{q} \vee \bar{s}q \vee \bar{r}ps.$

The result (1) itself is susceptible to yet a further simplification, covered by none of the seven. The initial clause of (1) is in fact redundant; (1) is equivalent to:

(2) $\bar{p}\bar{q} \vee \bar{s}q \vee \bar{r}ps.$

There is a quick way of testing any clause of an alternational normal schema to see if it can be thus dropped as redundant. The law (vii) of Chapter 9 tells us how: just check, by fell swoop, whether the clause implies the rest of the schema. The clause '$\bar{p}\bar{s}$' of (1) is found by fell swoop to imply the remainder of (1), and this marks '$\bar{p}\bar{s}$' as redundant in (1).

The schema (6) of the preceding chapter was already an alternational normal schema. (2) here is another, and, we see, equivalent to (6). Simplification can go a long way.

Sometimes an alternational normal schema can be simplified by dropping not a whole clause but just a literal. This happened already in the preceding chapter when we moved from (7) to (8) on the strength of (1) of Chapter 9. It can happen also in cases untouched by (1) of Chapter 9. An example is afforded by '$pq \vee p\bar{q}r \vee \bar{p}\bar{q}\bar{r}$', which proves equivalent to '$pq \vee pr \vee \bar{p}\bar{q}\bar{r}$'.

There is a quick way of testing a literal in some clause of an alternational normal schema to see if it can be dropped as redundant. Just see, by fell swoop, whether the rest of the clause implies the whole schema. Thus take '$pqr \vee p\bar{r} \vee \bar{q}\bar{r}$'. To check that the '$r$' of '$pqr$' is redundant here, we need merely check that 'pq' implies '$pqr \vee p\bar{r} \vee \bar{q}\bar{r}$'.

Why this method always works can be seen well enough by just analyzing the present example. We want to check that '$pqr \vee p\bar{r} \vee \bar{q}\bar{r}$' is equivalent to '$pq \vee p\bar{r} \vee \bar{q}\bar{r}$'. What we have checked by fell swoop is that 'pq' implies '$pqr \vee p\bar{r} \vee \bar{q}\bar{r}$', and this amounts to saying (by (vii), Chapter 9) that '$pqr \vee p\bar{r} \vee \bar{q}\bar{r}$' is equivalent to the combined alternation '$pq \vee pqr \vee p\bar{r} \vee \bar{q}\bar{r}$'. But then we reduce the first two clauses of this alternation to 'pq' (by (1), Chapter 9), getting '$pq \vee p\bar{r} \vee \bar{q}\bar{r}$'.

Sometimes the elimination of a redundant literal can reward us doubly, by engendering a clausal redundancy that we can drop as well. Thus consider again our last example, '$pqr \vee p\bar{r} \vee \bar{q}\bar{r}$'. We see by three fell swoops that none of its three clauses is redundant. After we drop the

redundant 'r', however, the remainder '$pq \vee p\bar{r} \vee \bar{q}\bar{r}$' has a redundant clause. Find it.

An alternational normal schema is quickly checked for consistency, as noted in the preceding chapter: each inconsistent clause drops as patently inconsistent. A validity check of alternational normal schemata is harder, but it is provided now by elimination of redundant literals. For, recall the redundancy test for a literal: it is that the rest of its clause implies the whole schema. But if the whole schema is in fact valid, anything implies it; so then each literal will fall away by the redundancy test until we are left with only one-letter clauses. This remainder will be not only valid but patently valid: something like '$p \vee \bar{q} \vee q$'.

Two good ways are now before us for simplifying alternational normal schemata. We can test a clause for redundancy, and we can test a literal for redundancy, in each case by fell swoop. An alternational normal schema can, however, resist both redundancy tests and still admit of simplification in more devious ways. An example is:

(3) $$p\bar{q} \vee \bar{p}q \vee q\bar{r} \vee \bar{q}r.$$

By twelve fell swoops the reader can test each clause and each literal of (3) for redundancy and draw a blank every time. Yet (3) has a simpler equivalent, '$p\bar{q} \vee \bar{p}r \vee q\bar{r}$'.

Simplifications by these fell-swoop techniques are always beneficial, even failing the assurance that we have found a shortest equivalent. We may forgo such further reductions as that of (3) to '$p\bar{q} \vee \bar{p}r \vee q\bar{r}$', or we may press on with the next paragraphs.

It is remarkable that no quick and general method is known for reducing an alternational normal schema to a shortest equivalent. We have to exhaust possibilities. An adequate range of possibilities can be staked out by the following considerations. An alternational normal schema is an alternation of clauses each of which *implies* the schema (since 'p' implies '$p \vee q$'). Further, if we have checked for redundant literals and deleted them, we may be sure that each clause is a *prime* implicant of the schema; that is, when you drop any literal from a clause, the remainder of the clause ceases to imply the schema. So, if we somehow assemble all the prime implicants of a given schema, we may be sure that we have caught all the clauses of any shortest alternational normal equivalent of our schema. At our clumsiest we may then simply try the various combina-

tions of prime implicants and select a shortest alternation that comes out equivalent to the original schema.

There is a mechanical method, due to Samson and Mills, for generating all the prime implicants of a schema. We start with an alternational normal schema that has been freed of redundant literals. This schema is an alternation of some of its prime implicants. Now if two of these clauses are opposed in one and only one letter, say 'p' (so that one clause has 'p' and the other has '$\bar{p}$'), take the conjunction of the rest of their literals. This conjunction (minus any duplications) is what I call the *consensus* of the two clauses. For instance, the first and third clauses of (3) have the consensus '$p\bar{r}$'; the second and fourth have '$\bar{p}r$'. It can be proved that this process of consensus-taking will turn up every missing prime implicant.[13]

Thus, to find a shortest normal equivalent of (3), we begin by generating its two missing prime implicants as above. Annexing them to (3), we have:

(4) $$p\bar{q} \vee \bar{p}q \vee q\bar{r} \vee \bar{q}r \vee p\bar{r} \vee \bar{p}r.$$

This is equivalent still to (3), by virtue of (vii) of Chapter 9; it is a redundant equivalent. We know that its last two clauses can be dropped; but if, instead, we keep one or both of them, perhaps we can drop others, and more. So we make various experimental fell swoops. In the course of them we find not only that '$\bar{p}q$' implies the rest of (4) and so can be dropped, but also that '$\bar{q}r$' implies the rest of what remains and so can be dropped in turn, and also that '$p\bar{r}$' implies the rest of that. We end up thus with '$p\bar{q} \vee \bar{p}r \vee q\bar{r}$', an improvement on (3). There is another equally short version that we could get too; the reader might like to search it out.

HISTORICAL NOTE: The problem of simplifying truth-functional schemata has interested industry, because of an application to electric circuits. Thus picture two terminals and two intervening switches. If the switches are connected in parallel, then the current is on just in case the one switch *or* the other is closed. If they are connected in series, then the current is on just in case the one switch *and* the other are closed. Such are the roles of alternation and conjunction; and as for negation, it answers to the throwing of a switch. There results, as Claude Shannon

[13] See my *Selected Logic Papers*, pp. 166 f.

observed in 1938, a correspondence between circuits and schemata. A practical technique for reducing a schema to a shortest equivalent schema would enable the engineer to reduce a circuit to a simplest equivalent circuit. I was still unaware of the elusiveness of such a technique as late as 1948, when I was working at the present book and hoping to base my whole treatment of truth-function logic upon an easy simplification routine. My first article on the problem appeared in 1952. By then the computation laboratory at Harvard had found it worthwhile actually to tabulate and publish the simplest equivalents of the 65,536 truth functions of four distinct letters. (See Aiken.) Because automata require complicated circuits, many studies of the simplification problem have been published in engineering media. It was in one such in 1954 that Samson and Mills presented the consensus method, as I am calling it, for getting all the prime implicants as in (4). In another such paper, in 1957, Ghazala showed a way of expediting the rest of the job, that of selecting an adequate minimum of prime implicants as in '$p\bar{q} \vee \bar{p}r \vee q\bar{r}$'. Meanwhile another engineer, Rolf K. Müller, showed me in 1955 that the major technical trouble lies in the surprising multitude of prime implicants, in the case of schemata with six or a dozen different letters. Fridshal cites a nine-letter schema which, he claims, has 1,698 prime implicants. Engineers have been programming computers to find simplifications of such schemata, but even so the task can be forbidding. A simplification technique that did not depend on any exhaustive survey of the prime implicants would be a boon.

EXERCISES

1. Check the equivalence of (1) to (2) by truth-value analysis.
2. Similarly for (3) and '$p\bar{q} \vee \bar{p}r \vee q\bar{r}$'.
3. Investigate these schemata for redundant clauses and redundant literals:

$$p\bar{q}\bar{r} \vee \bar{p}q \vee pr \vee qr, \qquad pq \vee p\bar{q}r \vee \bar{p}\bar{q}\bar{r}.$$

4. Do the same for the six alternational normal schemata obtained in Exercise 1 of Chapter 10.
5. Find another equivalent of (4), as short as '$p\bar{q} \vee \bar{p}r \vee q\bar{r}$'. Show steps.

6. Find all prime implicants of each of the six alternational normal schemata obtained in Exercise 1 of Chapter 10. (First drop any redundancies that were revealed just now by Exercise 4.)

7. Find a shortest alternational normal equivalent of each of the six.

12
DUALITY

All logical computation at the truth-functional level is essentially computation with '⊤' and '⊥'. Hence it is to be expected that two schemata will be quite parallel in their behavior if they are just alike under truth-value analysis except for a thoroughgoing interchange of '⊤' and '⊥'. Schemata so related are called *duals* of each other. They behave in relation to each other according to laws which, for their theoretical interest and occasional convenience, warrant some notice.

Though duals are opposed somewhat in the manner of '⊤' to '⊥', they are not to be confused with mere contradictories or mutual negations. The prime example of duality, rather, is conjunction versus alternation. Conjunction and alternation are alike except for a *thoroughgoing* interchange of '⊤' and '⊥', in the following sense. Conjunction, to begin with, is describable thus:

1ST COMPONENT	2ND COMPONENT	RESULT
⊤	⊤	⊤
⊥	⊤	⊥
⊤	⊥	⊥
⊥	⊥	⊥

Now to interchange '⊤' and '⊥' merely in the last column would indeed produce a truth function which is the negation of conjunction. Interchange '⊤' and '⊥' throughout all three columns, however, and what you get is precisely a description of alternation:

1ST COMPONENT	2ND COMPONENT	RESULT
⊥	⊥	⊥
⊤	⊥	⊤
⊥	⊤	⊤
⊤	⊤	⊤

Such is the sense in which 'pq' and '$p \vee q$' are said to be duals. In general the relationship between dual schemata S and S' is this: whenever each of 'p', 'q', etc. is interpreted oppositely for S and S', the truth values of S and S' turn out oppositely to each other.

Trivially, by this standard, '$\bar{p}$' is dual not to 'p' but to '$\bar{p}$' itself; for, give opposite values to 'p' and you get values for '$\bar{p}$' which are opposite to each other.

Instead of inverting '⊤' and '⊥' individually throughout the table, as was done above, we could get an equivalent tabulation of the dual by leaving the columns for components intact and just inverting the result column bodily (so that a top '⊤' becomes a bottom '⊥'). This can be verified in the above case, and it can be seen on reflection to work always. In particular then the mark of a self-dual truth function is that its result column looks the same upside down.

The duality of 'pq' to '$p \vee q$' is evident without resort to the above tabulation if we simply compare the original descriptions of conjunction and alternation. A conjunction is true when its components are all true, and otherwise false; whereas an alternation is false when its components are all false, and otherwise true. These two descriptions are alike except for interchange of 'true' and 'false'; hence 'pq' and '$p \vee q$' are bound to behave alike except for a thoroughgoing interchange of the roles of '⊤' and '⊥'—which is what duality means. The self-duality of '$\bar{p}$' is evident similarly from the general description of a negation as "true or false according as its component is false or true"; for, switch the words 'true' and 'false' in this description and you simply have the description of negation over again.

More generally now, consider any schema S built up of letters by means exclusively of negation, conjunction, and alternation (hence devoid of '$\rightarrow$' and '$\leftrightarrow$'). Suppose a second schema S' is like S except that it has alternation wherever S has conjunction and vice versa. Then truth-value analyses of S and S' are bound to match except for interchange of '⊤' and '⊥' throughout; for, we just saw that the explanations of conjunction and alternation are alike except for switching 'true' with 'false', and

that the explanation of negation is unchanged by switching 'true' with 'false'. So S and S' are duals.

What has just been established will be called the *first law of duality:* Where S is any truth-functional schema devoid of '$\rightarrow$' and '$\leftrightarrow$', the result of changing alternation to conjunction and vice versa throughout S is dual to S. This law affirms immediately the duality of 'pq' to '$p \vee q$', and the self-duality of '$\bar{p}$' and of 'p'. It affirms also the duality of '$\bar{p} \cdot q \vee \bar{r}$' to '$\bar{p} \vee q\bar{r}$', the duality of '$p\bar{q} \vee \bar{p}r$' to '$p \vee \bar{q} \cdot \bar{p} \vee r$', the duality of '$pq \vee qr \vee pr$' to '$p \vee q \cdot q \vee r \cdot p \vee r$', etc.

So we now have a quick and graphic way of forming a dual of a schema: interchange conjunction and alternation. This procedure depends, be it noted, on absence of '$\rightarrow$' and '$\leftrightarrow$'; but we can get rid of '$\rightarrow$' and '$\leftrightarrow$' in advance, since '$p \rightarrow q$' may be rendered as '$\bar{p} \vee q$' and '$p \leftrightarrow q$' as '$\bar{p} \vee q \cdot \bar{q} \vee p$' or '$pq \vee \bar{p}\bar{q}$'.

In interchanging conjunction and alternation to get duals, special care must be taken to preserve grouping. In case of doubt, think of full parentheses as restored in lieu of the dot conventions. Thus '$p \cdot q \vee r$' has as dual not '$p \vee q \cdot r$' but '$p \vee qr$'. For, '$p \cdot q \vee r$' means '$p(q \vee r)$', and '$p \vee qr$' means '$p \vee (qr)$', in which the same pattern of grouping is preserved; '$p \vee q \cdot r$', on the other hand, means '$(p \vee q)r$', and is dual rather to '$pq \vee r$'.

Given any schemata S and S', now, we can test whether S' is a dual of S by forming an explicit dual of S according to the above method and then checking it for equivalence to S'. In particular we can thus determine whether a given schema S is a dual of itself; we have merely to form the explicit dual of S by switching conjunction with alternation, as explained, and then to test this result for equivalence to S. Apart from trivial cases, such as '$\bar{p}$', self-duality is rather rare; but it does occur. E.g., '$pq \vee pr \vee qr$' is dual to itself, since it is equivalent to its own explicit dual '$p \vee q \cdot p \vee r \cdot q \vee r$'. (See Chapter 9, Exercise 4.)

Switching alternation with conjunction is not the only convenient way of forming a dual. Another way, which does not even require a preparatory elimination of '$\rightarrow$' and '$\leftrightarrow$', is provided by the *second law of duality:* If in any schema you negate all letters and also the whole, you get a dual. This law is evident from the original definition of duality; for negating the letters has the same effect as reversing all interpretations of letters, and negating the whole reverses the truth value of the outcome.

DeMorgan's laws themselves (Chapter 10) are essentially duality principles, as may be seen by rearguing them in the present context. As dual

of '$pq \ldots s$' the first law of duality cites '$p \vee q \vee \ldots \vee s$', whereas the second cites rather '$-(\bar{p}\bar{q} \ldots \bar{s})$'; these two must then be equivalent to each other, and thus DeMorgan's first law, (i) of Chapter 10, holds. (ii) admits of a parallel argument.

A *third law of duality* is this: A schema is valid if and only if its dual is inconsistent. For, if two truth-value analyses differ to the extent of a thoroughgoing interchange of '$\top$' and '$\bot$', clearly the one will show validity if and only if the other shows inconsistency.

Fourth law of duality: A schema S_1 implies a schema S_2 if and only if the dual of S_2 implies the dual of S_1. This is seen as follows. The duals of S_1 and S_2 behave like S_1 and S_2, under truth-value analysis, except for a switching of 'true' with 'false' throughout. Hence to say that no interpretation of letters makes S_1 true and S_2 false is the same as saying that no interpretation makes the dual of S_1 false and the dual of S_2 true.

Fifth law of duality: Schemata are equivalent if and only if their duals are equivalent. This follows from the fourth law, since equivalence is mutual implication.

The third, fourth, and fifth laws enable us, having established one validity or inconsistency or implication or equivalence by truth-value analysis or otherwise, to infer an additional inconsistency or validity or implication or equivalence without further analysis. E.g., having verified that '$p \vee \bar{q} \,.\, q \vee \bar{r} \,.\, r \vee s$' implies '$p \vee s$' (as may be done by the method of the fell swoop, Chapter 7), we may conclude by the fourth law of duality that 'ps' implies '$p\bar{q} \vee q\bar{r} \vee rs$'. This operation may, in contradistinction to the first, be spoken of as the *full swap*.

Either of DeMorgan's laws, (i)–(ii) of Chapter 10, follows from the other by the fifth law of duality. Again, from the law of distributivity of conjunction through alternation (Chapter 10) we can, by the fifth law of duality, infer a *law of distributivity of alternation through conjunction:*

'$p \vee qr \ldots t$' is equivalent to '$p \vee q \,.\, p \vee r \,.\ldots.\, p \vee t$'.

This law shows that conjunction and alternation are in still more congenial relations to each other than are multiplication and addition. In arithmetic we can multiply out, thus:

$$x(y + z + \ldots + w) = xy + xz + \ldots + xw,$$

but we cannot "add out" thus:

$$x + yz \ldots w = (x + y)(x + z) \ldots (x + w).$$

In the case of alternation and conjunction, on the other hand, distribution works both ways.

Indeed, since by the fifth law of duality all equivalences continue to hold when conjunction and alternation are switched, we may conclude at once that the technique of reducing a truth-functional schema to a normal schema may be reproduced entire with alternation and conjunction switched. This switched procedure issues in *conjunctional* normal schemata such as:

$$p \vee \bar{q} \vee r \,.\, \bar{p} \vee s \,.\, q \vee \bar{r} \vee s$$

—i.e, in conjunctions of alternations of literals. The *clauses* of a conjunctional normal schema are alternations, not conjunctions; and the schema is a conjunction of its clauses.

Because of duality, any procedure for simplifying alternational normal schemata has an exact parallel for conjunctional normal schemata. In particular, just as any inconsistent clause such as '$\bar{q}sq$' drops from an alternational normal schema (as long as some clause remains), so any valid clause such as '$\bar{q} \vee s \vee q$' drops from a conjunctional normal schema.

It was noted in Chapter 10 that the alternational normal form affords an immediate test of inconsistency. The conjunctional normal form affords a corresponding test of validity: just simplify by dropping valid clauses in the above fashion and see if nothing but a visibly valid clause remains.

Along with the developed alternational normal form that confronted us in Chapter 10, a developed conjunctional normal form is assured by duality. Here the appropriate operation of development is the one that carries 'p' into '$p \vee q \,.\, p \vee \bar{q}$'. The conjunctional normal form '$\bar{s} \vee q \,.\, \bar{r} \vee p \vee s$' has the developed conjunctional normal form:

$$p \vee q \vee r \vee \bar{s} \,.\, p \vee q \vee \bar{r} \vee \bar{s} \,.\, \bar{p} \vee q \vee r \vee \bar{s}$$
$$\,.\, \bar{p} \vee q \vee \bar{r} \vee \bar{s} \,.\, p \vee q \vee \bar{r} \vee s \,.\, p \vee \bar{q} \vee \bar{r} \vee s,$$

the dual of what we saw at the end of Chapter 10. For two developed conjunctional normal forms having the same letters, the test of implication is the reverse of what it was for developed alternational normal forms; all the clauses of the implied schema are among those of the implying schema. The mark of inconsistency in developed conjunctional

normal form is presence of all the 2^n possible clauses, and the mark of validity is disappearance.

A concluding word of caution: the alternational and conjunctional forms of a schema are ordinarily not duals of each other. Duals, after all, are not equivalent, except in the case of self-duality.

HISTORICAL NOTE: The essence of duality is DeMorgan's laws, which, as noted, go back six centuries to William of Ockham. More deliberate treatment of duality dates from Schröder (1877).

EXERCISES

1. Which of the schemata:

$$p \leftrightarrow q, \quad p \leftrightarrow \bar{q}, \quad \bar{p} \leftrightarrow q, \quad -(p \leftrightarrow q)$$

are dual to which? Justify.

2. Write the duals of these:

$$p \rightarrow q, \quad q \rightarrow p, \quad -(p \rightarrow q), \quad -(q \rightarrow p).$$

3. We saw in Chapter 11 a test of redundancy for a clause or a literal of an alternational normal schema. What, by considerations of duality, should be the tests of redundancy for a clause or a literal of a conjunctional normal schema?

4. By successive transformations, transform each of the schemata of Exercise 1 of Chapter 10 into a conjunctional normal schema. Simplify where you can.

5. Expand the results of Exercise 4 into developed conjunctional normal form.

6. Test the four schemata of Exercise 1 of Chapter 6 for validity by putting them into conjunctional normal form.

13 AXIOMS

The application of logic to one or another scientific theory, such as arithmetic or some branch of physics, is sometimes made explicit by setting up what is called an axiom system. Certain statements of the theory are chosen as a starting point under the name of axioms, and then further statements, called theorems, are generated by showing that they are logically implied by the axioms. The implication concerned here goes beyond the kind of implication that we have thus far considered, namely truth-functional implication, and exploits also the further resources of logic that we shall turn to in later parts of the book.

A variant of the axiomatic method has often been used within logic itself, and even within truth-function logic, to generate valid schemata. Here the generating relation can no longer be cited simply as implication, since implication here is ubiquitous; any valid schema is implied by any and every schema. Instead, specific formal rules of inference are given. A usual one is *modus ponens:* if a theorem (or axiom; axioms count as theorems) is a conditional whose antecedent is also a theorem, then put down the consequent as a theorem. Another usual one is *substitution:* substitute any schema for all occurrences of a letter in a theorem. If we start with valid axioms, obviously these two rules will lead only to valid theorems.

One interesting choice of axioms, due to Łukasiewicz, comprises these three:

(1) $p \to q \mathrel{.\to:} q \to r \mathrel{.\to.} p \to r,$

(2) $p \to .\, \bar{p} \to q,$

(3) $\bar{p} \to p \mathrel{.\to} p.$

Let us generate some theorems by the two rules. Substitution of 'p' for 'q' in (2) gives:

(4) $p \to .\, \bar{p} \to p.$

Substitution of 'p' for 'r' in (1) gives:

(5) $$p \to q \mathrel{.\to:} q \to p \mathrel{.\to.} p \to p.$$

Substitution of '$\bar{p} \to p$' for 'q' in (5) gives:

(6) $$p \to .\, \bar{p} \to p \mathrel{:\to:} .\, \bar{p} \to p \mathrel{.\to} p \mathrel{:\to.} p \to p.$$

Modus ponens, applied to (6) and (4), gives:

(7) $$\bar{p} \to p \mathrel{.\to} p \mathrel{:\to.} p \to p.$$

Modus ponens, applied to (7) and (3), gives:

(8) $$p \to p.$$

This sequence of steps is called a *proof* of '$p \to p$'. In a condensed notation the whole proof can be put thus:

$$\text{By (1)}, \qquad [2\ (q/p) \to: 3 \to.]\, p \to p.$$

The numeral '3' here stands as an abbreviation of the schema (3), hence '$\bar{p} \to p \mathrel{.\to} p$'. The expression '2 (q/p)' stands for the schema (4), this being what (2) becomes when 'q' is supplanted by 'p'. Thus the whole string of symbols amounts to (6). The mention of (1) means that this whole schema (6) can be got from (1) by certain substitutions. A reader presented merely with this condensed proof would recover (6), compare its structure with that of (1), and discover for himself that '$\bar{p} \to p$' and 'p' were the required substitutes for 'q' and 'r' in (1). The square brackets, finally, indicate excision by modus ponens—twice over, in this case.

This system visibly treats only of negation and the conditional. However, any truth-functional schema can be translated into those terms, since '$p \vee q$' and 'pq' are equivalent to '$\bar{p} \to q$' and '$\bar{} (p \to \bar{q})$'. There is a proof that every valid truth-functional schema, or its translation, can be got from the axioms (1)–(3) by the two rules. The system is in this sense *complete*.

Axiomatists are naturally concerned that their axioms be *independent:* that none be derivable as a theorem from the rest, and hence dispensable. The independence of an axiom is neatly established if we can so reinterpret the symbols as to falsify that axiom while still preserving the validity of the other axioms and the soundness of the rules of inference. For instance, we can show that the axiom (3) is independent, in this system,

by reinterpreting the negation sign as simply yielding a falsehood wherever applied. We keep the interpretation of '→'. Hence the rules of modus ponens and substitution remain sound and (1) remains valid. Also (2) comes out valid; for if '$\bar{p}$' comes out always false, '$\bar{p} \rightarrow q$' comes out always true. But (3), thus reinterpreted, comes out false for false 'p'.

Again, to show (2) independent, reinterpret the negation sign as yielding a truth wherever applied.

Various other complete systems of independent axioms have been contrived for truth-function logic. Not all of them are based on negation and the conditional. Some use negation and alternation. One, due to Nicod and improved by Łukasiewicz, uses only the truth-functional connective '|', which, we saw (Chapter 2), suffices for expressing all the truth functions. This system has just one axiom:

$$p \mid . q \mid r : \mid : : s \mid . s \mid s : \mid : . s \mid q . \mid : p \mid s . \mid . p \mid s.$$

Its rules of inference are substitution and a variant of modus ponens which says that if a theorem has the form '$A \mid . B \mid C$' and has a theorem in the 'A' place, then the schema in the 'C' place is a theorem. If we define '$p \rightarrow q$' as '$p \mid . q \mid q$', the axiom above becomes less bewildering:

$$p \mid . q \mid r : \mid : . s \rightarrow s . \mid : s \mid q . \rightarrow . p \mid s.$$

Axiomatic logic is one thing; the application of full-blown logic to extralogical axioms is another. The contrast between them is the contrast between what Sheffer called *foundational* and *postfoundational* systems. Foundational systems have their autonomous rules of inference; postfoundational systems just infer their theorems from their axioms by logical implication, deferring to logic for the analysis and the technology of that relation.

Still, foundational systems can be fashioned also for extralogical subject matter. An example is the following axiomatic algebra of subtraction.[14] There are two axioms:

$$x = x - (y - y), \qquad x - (y - z) = z - (y - x).$$

Substitution is again one of the rules of inference, and a second rule of inference allows us to put the left side of a theorem for the right in any theorem. Substitution in the second axiom, e.g., gives the theorem:

[14] From my *Selected Logic Papers*, pp. 54–60.

$$z - (x - (y - y)) = (y - y) - (x - z),$$

and from this and the first axiom by means of the second rule of inference we get the theorem:

$$z - x = (y - y) - (x - z).$$

A queerer example: the second rule of inference allows us to put the left side 'x' of the first axiom for the right side in that axiom itself; we get the theorem '$x = x$'.

We can get more. The system can be proved complete, in this sense: every equation that is constructed of this subtraction notation and is valid, or true for all values of its variables, is derivable from the two axioms by the two rules. Moreover, the notation is stronger than it looks; $x + y$ can be expressed in it, as $x - ((y - y) - y)$.

Usually, however, axiomatizations of extralogical subjects are postfoundational. There is a good reason: mass production. Once for all we develop logical techniques for establishing implication; afterward we can use these for deriving theorems from axioms in postfoundational systems on any subject.

On the other hand an axiom system for logic is necessarily foundational, and I would in conclusion remark further that it is of dubious value—especially in the logic of truth functions. This domain, after all, enjoys the luxury of a *decision procedure* for validity—that is, a mechanical test. Truth-value analysis affords one such test of validity; truth tables afford another; transformation into conjunctional normal form affords a third. Thus blessed, we should be unwise to make practical use of the axiomatic method in this domain. It is inferior in that it affords no general way of reaching a verdict of invalidity; failure to discover a proof for a schema can mean either invalidity or mere bad luck.

Łukasiewicz swore by axiomatic truth-function logic as a training ground for axiomatic method in more demanding domains. I swear rather by the sufficiency, unto the day, of the evil thereof. Axiomatic logic, with its schemata and its specific rules of inference, is very unlike postfoundational axiom systems. Training in the latter, if wanted, is best sought as such. Its main component is training in the recognition or proof of implication, since implication is what relates postfoundational axioms to the theorems; and such training is a pervading purpose of this book. There will be some specific notice of postfoundational axioms in Chapters 42 and 48. Meanwhile I have included this brief account of axiomatic

truth-function logic partly for contrast and partly because of the prominence of the subject in earlier literature.

HISTORICAL NOTE: Frege, in 1879, was the first to axiomatize the logic of truth functions and to state formal rules of inference. Frege's system, like the Łukasiewicz system (1)–(3), was couched in negation and the conditional and proceeded by modus ponens and substitution; but the axioms were more cumbersome, and the rule of substitution was not made explicit. Whitehead and Russell couched their system, 1910, in negation and alternation. The Łukasiewicz system dates from 1929.[15] The Nicod-Łukasiewicz axiom dates from 1917 and 1931. The axiomatic approach continued to dominate truth-function logic through the thirties, and a score of systems are in print. The first proof of completeness of such a system is due to Post, 1921.[16]

EXERCISES

1. Prove '$\bar{p} \to p \;.\!\to.\; \bar{p} \to q$' from (1)–(3) using the condensed proof notation.
2. Show the independence of (2) in detail.
3. In the Nicod-Łukasiewicz system, prove:

$$t \mid\!: s \mid\!. s \mid s \;.\!:\mid\!:: p \mid\!. q \mid r \;:\mid t \;.\!:\mid\!:. p \mid\!. q \mid r \;:\mid t$$

(which says '$t \mid\!. s \to s \;:\!\to\!:. p \mid\!. q \mid r \;:\mid t$' if we define '$\to$' as before).

[15] See Tarski's *Logic, Semantics, Metamathematics*, p. 43.

[16] For a completeness proof shorter than most, see my *Selected Logic Papers*, pp. 159–163.

II
GENERAL TERMS AND QUANTIFIERS

14
CATEGORICAL STATEMENTS

There are many simple and logically sound inferences for which the foregoing techniques are inadequate. An example is this:

No philosophers are wicked,	*Schematically:*	No G are H,
Some Greeks are philosophers;		Some F are G;
Therefore some Greeks are not wicked.		$\therefore$ Some F are not H.

Note that 'F', 'G', and 'H' here stand not for statements, after the manner of 'p' and 'q' in Part I, but for common nouns—or, in logical parlance, *general terms*.[1] Whether these nouns be thought of as substantive or adjective is an insignificant question of phrasing. 'G' appears as a substantive in the above example, viz. 'philosophers', and 'H' as an adjective, 'wicked'; but we could rewrite the adjective as a substantive, 'wicked individuals', if we liked. In the same spirit we can even treat intransitive verbs as terms, in effect, thus reckoning 'Some fishes fly' as a case of 'Some F are G'; for the difference between 'Some fishes fly' and 'Some fishes are flying things' is purely notational. The nouns or verbs which figure as terms may also, of course, be complex phrases such as 'employed for ten years by Sunnyrinse', 'wear brass rings in their noses', etc. Whether terms be thought of as in the singular or the plural is also a logically insignificant question of phrasing; thus there is no need to distinguish between 'No philosopher is wicked' and 'No philosophers are wicked', nor between 'All philosophers are wise' and 'Every philosopher is wise'. There is no need even to distinguish between 'Some Greek is a philosopher' and 'Some Greeks are philosophers', provided that, as will be our practice here, we understand 'some' always to mean simply 'at least one'.

But, for all the latitude accorded to the concept of term, it remains clear that terms are never statements; and this is why the techniques of Part I are inadequate to the inference exhibited above. Part I dealt with the

[1] What are spoken of as general terms or simply as terms in the present pages may, in view of developments in Parts III–IV, be designated more accurately as *monadic general terms*.

structures of compound statements relative only to their component statements; statements remained the smallest units of analysis. It is only now, in Part II, that we embark upon the analysis of those component statements in turn into the still smaller parts, not statements at all but terms, of which they are composed. Logically sound inferences depend for their soundness on the structures of the statements concerned, but the relevant structures may be either the broad outward structures studied in Part I or the finer substructures to which we are now turning. The example above is one which depends on structures of the latter kind.

It is the peculiarity of a statement to be true or false. It is the peculiarity of a term, on the other hand, to be *true of* many objects, or one, or none, and false of the rest. The term 'Greek' is true of each Greek, and the term 'wicked' is true of each wicked individual, and nothing else. The term 'natural satellite of the earth' is true of each natural satellite of the earth and nothing else, hence true of but one object, the moon. The term 'centaur' is true of each centaur and nothing else, hence true of nothing at all, there being no centaurs.

In place of the clumsy phrase 'is true of' we may also say 'denotes', in the best sense of this rapidly deteriorating word. But I prefer here to resist the temptation of good usage. 'Denotes' is so current in the sense of 'designates', or 'names', that its use in connection, say, with the word 'wicked' would cause readers to look beyond the wicked people to some unique entity, a quality of wickedness or a class of the wicked, as named object. The phrase 'is true of' is less open to misunderstanding; clearly 'wicked' is true not of the quality of wickedness, nor of the class of wicked persons, but of each wicked person individually.

When we are minded to speak of classes, the class of all the objects of which a term is true may, in keeping with a long tradition, be called the *extension* of the term. The extension of 'wicked' is thus the class of wicked persons; the extension of 'natural satellite of the earth' is the class whose sole member is the moon; and the extension of 'centaur' is the empty class. Terms may be said to *have* extensions, just as statements have truth values; but there is no need to think of a term as somehow a name *of* its extension, any more than there is to think of a statement as a name of its truth value. The use of terms proceeds smoothly on the whole without assumption of any special category of abstract objects called classes. It is ordinarily sufficient to know that a given term is *true of* this and that individual and false of the other, without positing any single collective entity called the term's extension. Some terms of set theory, indeed, cannot have extensions, not even the empty one; see Chapter 48.

Four ways of joining terms pairwise into statements have been treated as fundamental throughout the logical tradition stemming from Aristotle: 'All *F* are *G*', 'No *F* are *G*', 'Some *F* are *G*', and 'Some *F* are not *G*'. Statements of these four forms were called *categorical*. The four forms were distinguished by special nomenclature and by code letters '**A**', '**E**', '**I**', and '**O**', as follows.

A (Universal affirmative):	All *F* are *G*
E (Universal negative):	No *F* are *G*
I (Particular affirmative):	Some *F* are *G*
O (Particular negative):	Some *F* are not *G*

The form **A**, 'All *F* are *G*', may also be phrased 'If anything is an *F*, it is a *G*'; thus it is recognizable as the "generalized conditional" which was touched on in (1)–(3) of Chapter 3. Many other phrasings of **A** also come readily to mind: '*F* are *G*', 'Each (Every, Any) *F* is a *G*', 'Whatever is an *F* is a *G*', '*F* are exclusively *G*', 'Only *G* are *F*'.

E likewise has many phrasings: 'No *F* is (are) *G*', 'Nothing is both an *F* and a *G*', 'Nothing that is an *F* is a *G*', and even 'There is (are) no *FG*' (e.g., 'There is no black swan'), '*FG* do not exist'.

Correspondingly for **I**: 'Some *F* is (are) *G*', 'Something is both an *F* and a *G*', 'Something that is an *F* is a *G*', 'There is an *FG*', 'There are *FG*', '*FG* exist'. **O**, of course, has similar variants.

Often the terms properly answering to '*F*' and '*G*' are not directly visible at all in ordinary phrasing of statements. They may be partially covered up by such usages as 'nowhere', 'anywhere', 'always', 'everyone', 'whoever', 'whenever', etc. Thus the statement 'I go nowhere by train that I can get to by plane' is properly analyzable as of the form **E**, 'No *F* are *G*', but here we must understand '*F* as representing 'places I go to by train' and '*G*' as representing 'places I can get to by plane'. The statement 'Everyone in the room speaks English' has the form **A**, 'All *F* are *G*', where '*F*' represents 'persons in this room' and '*G*' represents 'speakers of English'.

In this last example the restriction to persons implicit in 'everyone' is essential, since there will be nonpersons in the room that do not speak English. In such an example as 'Everyone who pays his dues receives the Bulletin', on the other hand, 'everyone' is used instead of 'everything' only because of a habit of language, and not because the speaker feels any need of hedging his statement against such absurd objects as subhuman payers of dues. It would be pedantic to construe '*F*' for this example as

'persons who pay their dues', and quite proper to construe it as 'payers of dues'.

In putting statements of ordinary language over into the forms **A**, **E**, **I**, and **O** we must be on the alert for irregularities of idiom, and look beneath them to the intended sense. One such irregularity is omission of '-ever', as in 'Who hesitates is lost', 'I want to go where you go', 'When it rains it pours', 'She gets what she goes after'. Another irregularity is the non-temporal use of 'always', 'whenever', 'sometimes', 'never'. E.g., the statement:

> The sum of the angles of a triangle is always equal to two right angles

really means:

> The sum of the angles of any triangle is equal to two right angles,

and may be rendered 'All *F* are *G*' where '*F*' represents 'sums of angles of triangles' and '*G*' represents 'equal to two right angles'.

Frequently an **I** construction having to do with time is implicit in the inflection of a verb; witness 'We saw Stromboli and it was erupting', which comes out as 'Some *F* are *G*' with '*F*' construed as 'times we saw Stromboli' and '*G*' as 'times Stromboli was erupting'. (See Chapter 8, (5).) Further examples of temporal idioms which call for a little reflection, if the logical structure is to be properly extracted, are:

> I knew him before he lost his fortune,
> I knew him while he was with Sunnyrinse.

'Before' and 'while' here appear in the guise of statement connectives, like 'and' or 'or' or '$\rightarrow$'. But the statements are better analyzed as of the form **I**, 'Some *F* are *G*', where '*F*' represents 'moments in which I knew him' and '*G*' represents in the one case 'moments before he lost his fortune' and in the other case 'moments in which he was with Sunnyrinse'.

Reflection, indeed, should be the rule. Proper interpretation is not generally to be achieved through slavish dependence upon a check-list of idioms. 'Always' usually means 'at all moments', but it would be unjust to construe 'Tai always eats with chopsticks' as 'Tai eats with chopsticks at all moments'. The proper interpretation of this example is 'All *F* are *G*'

where '*F*' represents 'moments at which Tai eats' (not simply 'moments') and '*G*' represents 'moments at which Tai eats with chopsticks'.

The importance of reflecting upon context and the common sense of the concrete situation, rather than looking to any mere glossary, is manifest even in so basic a construction as 'An *F* is *G*'. 'A lady is present' is surely of the form **I**, but 'A Scout is reverent' is more likely to be intended in the form **A**. Caution is similarly needed in equating 'any' with 'every'; for, whereas the statements:

John can outrun every man on the team,
John can outrun any man on the team

need no distinguishing, a divergence appears as soon as 'not' is applied:

John cannot outrun every man on the team,
John cannot outrun any man on the team.

The first two statements are indistinguishably 'All *F* are *G*' (where '*F*' is 'man on the team' and '*G*' is 'whom John can outrun'); the third, however, is 'Some *F* are not *G*', while the fourth is 'No *F* are *G*'.

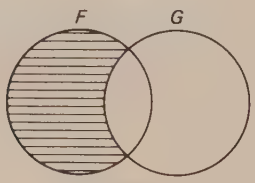
EXERCISE

Classify the following statements as between **A**, **E**, **I**, and **O**, and specify in each case what terms answer to '*F*' and '*G*'.

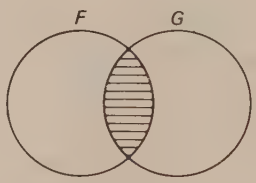
Blessed are the meek.
All that glisters is not gold.
There is no god but Allah.
Hope springs eternal.
The rule applies to everyone.
We should all be as happy as kings.
A policeman's lot is not an 'appy one.
There are smiles that make you blue.
I journeyed hither a Boeotian road.
I was stopped by the door of a tomb.

15
VENN'S DIAGRAMS

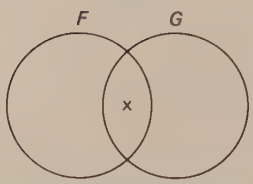
In Venn's diagrammatic method (1880), overlapping circles are used to represent the two terms of a categorical statement. The region in which the two circles overlap represents the objects which are both F and G. This region, called a *lens* in geometry, is shaded in Diagram 2. Where 'F' is taken as 'French' and 'G' as 'generals', this region represents the French generals. Correspondingly the part of the F-circle which lies outside the G-circle represents the objects which are F but not G: the French nongenerals, in the example. This region, called a *lune* in geometry, is shaded in Diagram 1. The significance of shading is emptiness; thus Diagram 2 affirms that no F are G, while Diagram 1 affirms that no F are other than G, or in other words that all F are G.



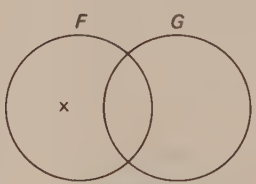
A: All *F* are *G*
Diagram 1



E: No *F* are *G*
Diagram 2



I: Some *F* are *G*
Diagram 3



O: Some *F* are not *G*
Diagram 4

Whiteness of a region in a Venn diagram means nothing but lack of information. In Diagram 2 the two lunes are left unshaded not because we think there are F which are not G and G which are not F, but because 'No F are G' gives us no information on the subject. All that 'No F are G' says is that the lens is empty, and this is all the information that Diagram 2 records. Similarly the lens and the right-hand lune in Diagram 1 are left

unshaded merely because 'All F are G' gives us no information concerning these further regions.

The great region outside both circles represents the objects, if any, which are neither F nor G. It is left blank in Diagrams 1 and 2 because 'All F are G' throws no light on such objects, and neither does 'No F are G'.

So, whereas shading means emptiness, nonshading does not assure nonemptiness. For nonemptiness another symbol is used, viz., a cross. Thus 'Some F are G', which affirms nonemptiness of the lens, is expressed by putting a cross in the lens as in Diagram 3. Here again the blankness of the other areas implies neither emptiness nor nonemptiness, but represents mere lack of information.

'Some F are not G', finally, affirms no more nor less than that the part of the F-circle which lies outside the G-circle has something in it; so it is represented by putting a cross in that lune as in Diagram 4.

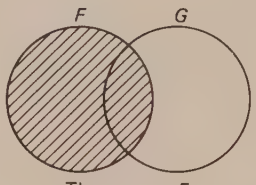
Certain simple laws of categorical statements are graphically reflected in the diagrams. The symmetry of Diagram 2, and of Diagram 3, reflects the fact that in **E** and **I** the order of terms is inessential: 'No F are G' amounts to 'No G are F', and 'Some F are G' to 'Some G are F'. Such switching of terms was known traditionally as *simple conversion*. The lopsidedness of Diagrams 1 and 4 reflects the fact that simple conversion is not in general applicable to **A** or **O**: 'All Greeks are men' is not to be confused with 'All men are Greeks', nor 'Some men are not Greeks' with 'Some Greeks are not men'.

A and **O** are mutual *contradictories*, or negations: **A** is true if and only if **O** is false. This relationship is reflected in the diagrams by the fact that Diagram 1 shows shading where, and only where, Diagram 4 shows a cross. With respect to blankness, or lack of information, Diagrams 1 and 4 are alike; with respect to information they simply and directly deny each other. Similarly **E** and **I** are mutual contradictories: **E** is true if and only if **I** is false.

A and **E**, 'All F are G' and 'No F are G', may also be felt to be somehow opposite to each other; however, their opposition is no matter of mutual negation, for we cannot say in general that **A** is true if and only if **E** is false. On the contrary, examples chosen at random are as likely as not to cause **A** and **E** to come out *both* false; this happens in particular when 'F' is taken as 'French' and 'G' as 'generals'. Similarly **I** and **O** are quite commonly both true, of course, as in this same example. But **A** and **O** are never both true nor both false, and similarly for **E** and **I**; here are the pairs of contradictories, or mutual negations.

74300

Whereas **A** and **E** are very commonly both false and **I** and **O** are very commonly both true, it is less common for **I** and **O** to come out both false, or for **A** and **E** to come out both true; but these things will happen where there are no *F*. Clearly, where there are no *F*, 'Some *F* are *G*' and 'Some *F* are not *G*' will both be false. Also, where there are no *F*, 'No *F* are *G*' will obviously be true; and yet 'All *F* are *G*' will likewise be true, in that there will be no *F* which is not *G*. These points are brought out diagrammatically by shading the *F*-circle in its entirety, as in Diagram 5, to mean that there are no *F*. This diagram verifies both **A** and **E**, for it shows both of the areas shaded which are shaded in Diagrams 1 and 2; and it falsifies both **I** and **O**, for it shows shading in place of both crosses of Diagrams 3 and 4.



There are no *F*
Diagram 5

A, 'All *F* are *G*', would seem at first glance to be stronger than **I**, 'Some *F* are *G*', and to imply it; but it does not, because of the possibility of there being no *F*. Diagram 5 depicts the very situation where, though **A** holds, **I** fails. It may happen that all my dimes are shiny (in that I have no dime to the contrary), and yet be false that some of my dimes are shiny, simply because I have no dimes at all. The most we can say is that if all *F* are *G and there are F* then some *F* are *G*.

If the reader thinks it odd to say that all of one's dimes are shiny when one has no dimes, he is perhaps interpreting 'All *F* are *G*' to mean, not simply 'There is no *F* that is not *G*', but 'There are *F* and each of them is *G*'. This, however, even if it be one of several defensible interpretations of an ambiguous idiom, is clearly not the interpretation which would make **A** the simple contradictory, or negation, of **O**: 'Some *F* are not *G*'. It is the general logical practice, and a convenient one, to understand 'All *F* are *G*' simply as the contradictory of **O**.

Diagram 5 as it stands was seen to reflect the fact that **I** does not follow from **A**. But Venn diagrams can also be used for constructive ends, as in showing that **I** follows from **A** supplemented with 'There are *F*'. To show

this we set down the diagram for **A**, viz. Diagram 1, and then enter 'There are F' into the diagram by putting a cross in the F-circle. The cross must go in the unshaded part of the F-circle, since the shaded part is known to be empty. So the result, showing a cross in the lens as it does, verifies **I**.

What has been said of the relationship between **A** and **I** applies equally to **E** and **O**: from **E**, 'No F are G', we may infer **O**, 'Some F are not G', only if we make the further assumption that there are F. Diagram 5 shows the situation where, though **E** holds, **O** fails. But we can show that **O** follows from **E** and 'There are F', by putting a cross in the F-circle of Diagram 2 and observing that we have verified **O**.

Finally let us observe a couple of simple inferences in which a conclusion is drawn from just a single premise, or assumption, instead of from two:

Some F are G,	There are no F,
$\therefore$ There are G.	$\therefore$ No F are G.

These inferences are justified respectively by Diagrams 3 and 5; for Diagram 3 shows a cross in the G-circle in support of the conclusion 'There are G', and Diagram 5 shows a shaded lens in support of the conclusion 'No F are G'.

EXERCISES

1. Does **A**, **E**, **I** or **O** follow from 'There are no G'? Does **A**, **E**, **I**, or **O** conflict with 'There are no G'? Appeal to diagrams.
2. Make a diagram for 'All F are G and all G are F'. Is this compatible with 'No F are G'? Explain.

16
SYLLOGISMS

What are spoken of traditionally as *syllogisms*[2] are arguments wherein a categorical statement is derived as conclusion from two categorical statements as premises, the three statements being so related that there are altogether just three terms, each of which appears in two of the statements. Six examples follow; one of them was already noted in Chapter 14.

All men are mortal, All Greeks are men; ∴ All Greeks are mortal.	All G are H, All F are G; ∴ All F are H.
No men are perfect, All Greeks are men; ∴ No Greeks are perfect.	No G are H, All F are G; ∴ No F are H.
All philosophers are wise, Some Greeks are philosophers; ∴ Some Greeks are wise.	All G are H, Some F are G; ∴ Some F are H.
No philosophers are wicked, Some Greeks are philosophers; ∴ Some Greeks are not wicked.	No G are H, Some F are G; ∴ Some F are not H.
All Greeks are men, Some mortals are not men; ∴ Some mortals are not Greeks.	All H are G, Some F are not G; ∴ Some F are not H.
Some men are not Greeks, All men are mortal; ∴ Some mortals are not Greeks.	Some G are not H, All G are F; ∴ Some F are not H.

A "valid" syllogism, ordinarily so-called, is a syllogism of such form as to be incapable of leading from true premises to a false conclusion. An

[2] *Categorical* syllogisms, more specifically, to distinguish them from *hypothetical* syllogisms, which are certain truth-functional arguments manageable by the methods of Part I.

easy test of validity of syllogisms is afforded by Venn's diagrams. Three overlapping circles are used, as in Diagrams 6 and 7, to represent the three terms '*F*', '*G*', and '*H*' of the syllogism. We inscribe the content of

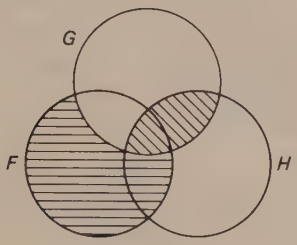


Diagram 6

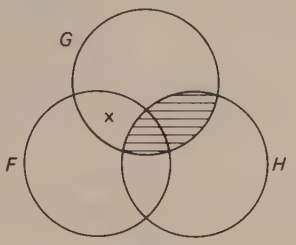


Diagram 7

the two premises into the diagram by the method explained in connection with Diagrams 1–4, and then we inspect the diagram to see whether the content of the conclusion has automatically appeared in the diagram as a result. Thus, let us test the second syllogism of the list above. We record its first premise, 'No *G* are *H*', by shading the lens common to the *G*-circle and the *H*-circle; then we record the second premise, 'All *F* are *G*', by shading the lune which lies in the *F*-circle outside the *G*-circle. The result is Diagram 6. It bears out the desired conclusion 'No *F* are *H*', since the lens common to the *F*-circle and the *H*-circle is fully shaded.

Let us next test the fourth of our six examples. We record the first premise, 'No *G* are *H*', as before, and then we record the second premise, 'Some *F* are *G*', by putting a cross in what remains of the lens common to the *F*-circle and the *G*-circle; the result is Diagram 7. It bears out the conclusion 'Some *F* are not *H*', there being a cross in the *F*-circle outside the *H*-circle.

It is left to the reader to construct diagrams verifying the remaining four of the above six syllogisms. (In the last of them, the second premise should be handled first; note why.)

The diagrammatic method can be used to determine not merely whether a given conclusion follows from given premises, but whether any conclusion at all (of a syllogistic kind) is capable of following from given premises. For, the conclusion—in order to be the conclusion of a so-called syllogism at all—must be 'All *F* are *H*', 'No *F* are *H*', 'Some *F* are *H*', or 'Some *F* are not *H*'; hence, unless the two bottom circles of the finished diagram exhibit one of the four patterns shown in Diagrams 1–4, there is no conclusion. In Diagram 8, e.g., the bottom circles exhibit none of the four patterns of Diagrams 1–4; and this shows that the premises

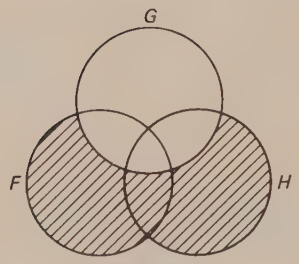


Diagram 8

'All H are G' and 'All F are G' cannot be the premises of any valid syllogism at all.

As a further example consider the premises 'All G are H' and 'All G are F'. These are recorded in Diagram 9; and we see that the diagram justifies no categorical conclusion in 'F' and 'H'. But this pair of premises is interesting in that it *almost* justifies a categorical conclusion in

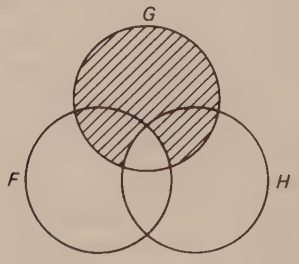


Diagram 9

'F' and 'H', viz., 'Some F are H'. If we add just the further premise 'There are G', to allow us to put a cross in the one part of the G-circle that remains unshaded, we then find the conclusion 'Some F are H' justified by a cross common to the F-circle and the H-circle. Thus the reinforced syllogism:

All Spartans are brave,	All G are H,
All Spartans are Greeks,	All G are F,
There are Spartans;	There are G;
$\therefore$ Some Greeks are brave	$\therefore$ Some F are H

is valid.

In the traditional terminology the term which plays the role of 'F' in 'All F are G', 'No F are G', 'Some F are G', or 'Some F are not G' is

called the *subject* of the statement. The other term, playing the role of '*G*', is called the *predicate*. The predicate of the conclusion is called the *major term* of the syllogism, and the subject of the conclusion is called the *minor term* of the syllogism. The remaining term, occurring in both premises but not in the conclusion, is called the *middle term* of the syllogism. Thus all the foregoing examples have been lettered in such a way as to make '*F*' the minor term, '*G*' the middle term, and '*H*' the major term.

The premise which contains the middle amd major terms is called the *major premise* of the syllogism. The other premise, containing the middle and minor terms, is called the *minor premise*. Thus all the foregoing examples have been stated with the major premise first and the minor premise second.

Medieval logicians had a scheme for coding the various forms of syllogisms. They stipulated the respective forms of the premises and conclusion (as among **A**, **E**, **I**, and **O**) by a triple of letters; thus '**EAO**' meant that the major premise was of form **E**, the minor premise **A**, and the conclusion **O**. This much was said to indicate the *mood* of a syllogism. But, even given the mood of a syllogism, there remains the question whether the major premise has the major term as subject and the middle term as predicate, or vice versa; and correspondingly for the minor premise. The four possibilities of arrangement which thus arise are called *figures,* and referred to by number as follows:

	1ST	2ND	3RD	4TH
Major premise:	*GH*	*HG*	*GH*	*HG*
Minor premise:	*FG*	*FG*	*GF*	*GF*
Conclusion:	*FH*	*FH*	*FH*	*FH*

Specification of mood and figure determines the form of a syllogism completely. Thus the six examples at the beginning of the present chapter are respectively **AAA** in the first figure, **EAE** in the first figure, **AII** in the first figure, **EIO** in the first figure, **AOO** in the second figure, and **OAO** in the third figure.

The fourth of the six above, viz. **EIO** in the first figure, can be given variant forms by simple conversion (see preceding chapter) of one or both premises. We thus get:

No *H* are *G*,
Some *F* are *G*;
∴ Some *F* are not *H*.

No *G* are *H*,
Some *G* are *F*;
∴ Some *F* are not *H*.

No *H* are *G*,
Some *G* are *F*;
∴ Some *F* are not *H*.

These are **EIO** in the second, third, and fourth figures. Similarly **EAE** and **AII** in the first figure (the second and third of the examples at the beginning of the chapter) are carried by simple conversion into **EAE** in the second figure and **AII** in the third.

The four syllogisms last mentioned, viz. **EAE** in the first and second figures and **AII** in the first and third, can be carried over into four further syllogisms by simple conversion of each of their conclusions. If we do this, though, we must afterward reletter '*F*' as '*H*' and '*H*' as '*F*' throughout the results in order that '*F*' may continue to represent the minor term and '*H*' the major term; also we must switch the order of the premises, so that the major premise may continue to appear first. The results, which the reader will do well to reproduce, are **AEE** in the fourth and second figures and **IAI** in the fourth and third.

Altogether, then, we have found fifteen valid forms:

FIRST FIGURE	SECOND FIGURE
AAA, EAE, AII, EIO	**EAE, AEE, EIO, AOO**
THIRD FIGURE	FOURTH FIGURE
IAI, AII, OAO, EIO	**AEE, IAI, EIO**

Note that no two of these fifteen have the same premises, when differences of figure are taken into account. We have here fifteen different pairs of premises, each with its appropriate conclusion. And it is readily verified by inspection of diagrams that none of these fifteen pairs of premises justifies any further syllogistic conclusion in addition to the one here indicated for it.

Viewed in terms merely of combinations and without regard to the existence of a valid conclusion, there are sixty-four possibilities for the premises of a syllogism. They may be **AA**, or **AE**, or **AI**, or **AO**, or **EA**, or **EE**, etc., to sixteen possibilities, and each of these sixteen may occur in any of four figures. In addition to the fifteen pairs of premises which have been found to yield valid syllogisms, therefore, there are forty-nine further pairs to consider. Now we saw in connection with Diagrams 8 and 9 how to check whether a given pair of premises justifies any syllogistic conclusion at all. If the reader so tests these forty-nine pairs (an hour's pastime), he will find that none of them justifies a syllogistic conclusion. The fifteen forms of syllogism listed above are the only valid ones.

In addition, however, nine forms come in for honorable mention.

These nine are forms which, like the above example of the Spartans, need a small reinforcing premise. 'There are F' fills the bill for five of them, 'There are G' for three, and 'There are H' for one. Let me simply record the nine in tabular fashion:

FIRST FIGURE	SECOND FIGURE	THIRD FIGURE	FOURTH FIGURE	ADDED PREMISE
AAI, EAO	**AEO, EAO**		**AEO**	There are F
		AAI, EAO	**EAO**	There are G
			AAI	There are H

Inferences involving so-called singular statements such as 'Socrates is a man', e.g.:

All men are mortal, Socrates is a man; ∴ Socrates is mortal,

were traditionally fitted into the syllogistic mold by treating the singular statements as of the form **A**. This procedure is artificial but not incorrect; we can construe 'Socrates is a man' as 'All G are H' where 'G' represents 'things identical with Socrates'. The above inference thus was classified as **AAA** in the first figure. But we shall end up, in Part IV, with a different treatment of singular inference.

In traditional logic it was customary to propound various rules whereby to test the validity of a syllogism. Examples: every valid syllogism has a universal premise (**A** or **E**); every valid syllogism has an affirmative premise (**A** or **I**); every valid syllogism with a particular premise (**I** or **O**) has a particular conclusion; every valid syllogism with a negative premise (**E** or **O**) has a negative conclusion. There are further rules whose formulation depends on a concept of "distribution" which has been omitted from the present exposition. As a practical method of appraising syllogisms, rules are less convenient than the method of diagrams. Indeed, the very notions of syllogism and mood and figure need never have been touched on in these pages, except out of consideration for their prominence in logic during two thousand years; for we can apply the diagram test to a given argument out of hand, without pausing to consider where the argument may fit in the taxonomy of syllogisms. The diagram test is equally available for many arguments which do not fit any of the arbitrarily delimited set of forms known as syllogisms.

HISTORICAL NOTE: The terms 'categorical' and 'syllogism' derive from Aristotle. So do the fourfold classification of categorical statements, and the nomenclature of the parts of a syllogism, and the classification into moods and figures, except that the fourth figure is attributed rather to his pupil Theophrastus. The notion of distribution and the rules that involve it are medieval.

EXERCISES

1. Construct diagrams verifying the remaining four of the six syllogisms at the beginning of the chapter.

2. Determine by diagrams what syllogistic conclusion, if any, follows from each of the following pairs of premises.

   All who blaspheme are wicked; No saint blasphemes.
   No snakes fly; Some snakes lay eggs.
   Nothing that lays eggs has feathers; Some fishes have feathers.
   Whatever interests me bores George; Whatever interests Mabel bores George.
   Whatever interests me bores George; Whatever interests George bores Mabel.

3. For each of the pairs in Exercise 2 which failed to yield a syllogistic conclusion, determine by diagram whether a supplementary premise of the form 'There are F', or 'There are G', or 'There are H', would suffice to bring forth a syllogistic conclusion.

17
LIMITS OF THESE METHODS

The inferences to which we have thus far been applying Venn's diagrams have all been made up of categoricals **A**, **E**, **I**, or **O** plus an occasional auxiliary of the type 'There are *F*'. Actually the diagrams can be used somewhat more widely; e.g., in arguing from the

PREMISES: Everyone east of the tracks is either slovenly or poor,
Not everyone east of the tracks is poor

to the

CONCLUSION: Some slovenly persons are not poor.

We set up a three-circle diagram as usual, wherein '*F*' means 'slovenly persons', '*G*' means 'persons east of the tracks', and '*H*' means 'poor persons'. Then the first premise is entered in the diagram by shading as empty just that compartment of the circle *G* which lies outside both *F* and *H*; see Diagram 10. The compartment thus shaded is neither a lune nor a lens, but a third shape. It represents persons east of the tracks who are neither slovenly nor poor; and just such persons are denied existence by the first premise. Now the second premise, which says in effect 'Some *G* are not *H*', is recorded as usual by putting a cross in what remains of *G* outside *H*. The result is seen to substantiate the conclusion 'Some *F* are not *H*', since there is a cross in *F* outside *H*.

An innovation due to C. I. Lewis (1918) is the use of a long bar instead of the cross in Venn's diagrams. The advantage of the bar is that it can be

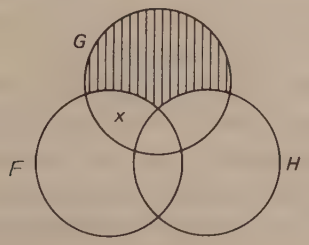


Diagram 10

made to lie across a boundary and thus indicate nonemptiness of a compound region. This innovation is useful in reasoning, e.g., from the

PREMISES: All of the witnesses who hold stock in the firm are employees,
Some of the witnesses are employees or hold stock in the firm

to the

CONCLUSION: Some of the witnesses are employees.

We set up a three-circle diagram in which '*F*' means 'witnesses', '*G*' means 'stockholders in the firm', and '*H*' means 'employees'. The lens common to *F* and *G*, then, stands for the witnesses who hold stock in the firm; so, on the basis of the first premise, we shade the part of that lens which lies outside *H*. (See Diagram 11.) Next, on the basis of the other premise, we run a bar through as much of the unshaded *F* as lies within *H* or *G*. The meaning of the bar is that one or another part of the total region marked by the bar has something in it. But the bar lies wholly within *F* and *H*; so the conclusion is sustained.

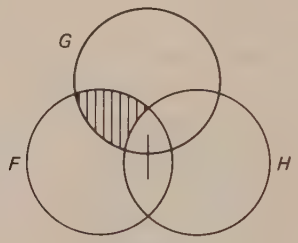


Diagram 11

The utility and versatility of Venn's diagrams are particularly evident from these last two examples. A shortcoming of the diagrams, however, is that they lend themselves less readily to arguments involving four or more terms. A diagram of overlapping ellipses can be constructed for four-term arguments, but it calls for careful drawing; and there is no way of constructing a diagram which will exhibit *five* or more reasonably simple regions in all combinations of overlapping. Where many terms are involved we may, however, try to break the argument down into parts involving manageably few terms. The following example is from Lewis Carroll:

PREMISES:
(1) The only animals in this house are cats;
(2) Every animal is suitable for a pet, that loves to gaze at the moon;
(3) When I detest an animal, I avoid it;
(4) No animals are carnivorous, unless they prowl at night;
(5) No cat fails to kill mice;
(6) No animals ever take to me, except what are in this house;
(7) Kangaroos are not suitable for pets;
(8) None but carnivora kill mice;
(9) I detest animals that do not take to me;
(10) Animals that prowl at night always love to gaze at the moon.

CONCLUSION: I always avoid a kangaroo.

This argument can be broken down as follows. From (1) and (5) we get the *lemma* or intermediate conclusion:

(11) All animals in this house kill mice;

this is the sort of step to which a simple three-term diagram is adequate. From (8) and (11), similarly, we get:

(12) All animals in this house are carnivora.

From (4) and (12) we get:

(13) All animals in this house prowl at night.

From (6) and (13) we get:

(14) All animals that take to me prowl at night.

Step by step in this fashion we can proceed to our desired conclusion, never using more than a three-term diagram for any one step. (If the reader cares to carry this through in detail, he should think of the universe as limited for purposes of the argument to animals—thus never bothering with 'animals' itself as a term.)

So we see that the purely mechanical method of diagrams proves inadequate when an argument turns on a *large number of terms;* a supplementary technique has to be invoked, such as that of resolving the argument into parts. Now another place where the unaided method of diagrams bogs down is where there is an *admixture of truth functions,* as in the following example:

PREMISES: If all applicants who received the second announcement are of the class of '00, then some applicants did not receive the second announcement.

Either all applicants received the second announcement or all applicants are of the class of '00.

CONCLUSION: If all applicants of the class of '00 received the second announcement then some applicants not of the class of '00 received the second announcement.

If we assign '*F*', '*G*', and '*H*' in obvious fashion, the inference takes on the form:

All *F* who are *G* are *H* → some *F* are not *G*,
All *F* are *G* ∨ all *F* are *H*,
∴ All *F* who are *H* are *G* → some *F* who are not *H* are *G*.

Diagrams are suited to handling the components 'All *F* who are *G* are *H*', 'Some *F* are not *G*', etc., and the methods of Part I are suited to '→' and '∨'; but just how may we splice the two techniques in order to handle a combined inference of the above kind?

So it is time to address ourselves to a more comprehensive theory. The three formulas last set forth, containing '*F*', '*G*', and '*H*', are *schemata* of a sort, but differ from the schemata of Part I in containing '*F*', '*G*', etc. and such words as 'all', 'some', 'who are', etc., to the exclusion of '*p*', '*q*', etc. In the ensuing chapters such schemata will be reduced to a logical notation subject to a ''decision procedure''—i.e., a mechanical routine for deciding validity, implication, consistency, etc. Such a decision procedure exists for truth-functional schemata in truth-value analysis; for the new class of schemata, however, the procedure will have to be more elaborate. Once it is at hand, all inferences of the sort we have been considering in the present Part—including the stubborn last example—can be adjudged mechanically by an implication test on premises and conclusion.

EXERCISES

1. Check the soundness of each of these inferences by diagram:

   All of the witnesses who hold stock in the firm are employees,
   All of the witnesses are employees or hold stock in the firm;
   ∴ All of the witnesses are employees.

   Everyone who knows both George and Mabel admires Mabel.
   Some who know Mabel do not admire her.
   ∴ Some who know Mabel do not know George.

   No European swans are black.
   All black swans are European.
   ∴ No swans are black.

   Only high-school graduates who can read French are eligible.
   Some can read French who are not high-school graduates.
   ∴ Some who are not eligible can read French.

   Everything is either a substance or an attribute.
   Modes are not substances.
   ∴ Modes are attributes

   Hint: Be prepared to shade the limitless region outside all circles.

2. Finish the kangaroo argument and supply a three-term diagram to justify each step.

3. Using '*F*', '*G*', and '*H*' in the obvious way, make a Venn diagram for each of the following two statements (Ullian):

   A thing is flexible if and only if it is either granulated or heavy.
   If nothing flexible is heavy then not everything granulated is heavy.

18
BOOLEAN SCHEMATA

To say of an object x that it is an F, we write 'Fx'. Here then is a new sort of sentence schema: 'Fx', 'Gx', etc. These may be compounded by truth functions; e.g.

(1) $$Fx \,.\, -Gx \,.\vee.\, Gx \rightarrow Hx.$$

We may conveniently abbreviate such compounds by extracting the 'x' everywhere and putting it at the end, thus:

$$[F\bar{G} \vee (G \rightarrow H)]x.$$

We arrive in this way at schematic representations of certain complex terms: schemata such as '$\bar{G}$', '$F\bar{G}$', '$G \rightarrow H$', '$F\bar{G} \vee (G \rightarrow H)$', etc. They will be called *Boolean term schemata*.

Where 'F' is interpreted as 'black' and 'G' as 'swan', '$\bar{G}$' is 'nonswan'; 'FG' is 'black swan'; '$F\bar{G}$' is true of just the other black things; and '$F \vee G$' is true of swans and black things generally. '$F \rightarrow G$', 'if black then a swan', is true of everything that is a swan if black—hence all swans and all that is not black; it is equivalent to '$\bar{F} \vee G$', or '$-(F\bar{G})$'.

When 'and' and 'or' are applied to terms, there is a curious oscillation to allow for. The things that are swans *or* black ($F \vee G$) are the swans *and* the black things. The things that are swans *and* black (FG) are not the swans *and* the black things, but just the black swans.

The concepts of validity, consistency, and implication carry over naturally to Boolean term schemata from the old truth-functional schemata. A term schema is called valid if, when the 'x' is restored, the resulting sentence schema is truth-functionally valid in the old sense. E.g., since '$Fx \vee -Fx$' or '$(F \vee \bar{F})x$' is valid, we call '$F \vee \bar{F}$' itself valid. Correspondingly for consistency and implication; thus '$F\bar{F}$' is called inconsistent and 'FG' is said to imply '$F \vee G$'.

The significance of such validity on the part of a term schema is that every interpretation of its letters turns it into a term that is true of every object x. But interpretation in what sense? In talking of interpretations of

term letters, we do best to allow freedom in choosing the *universe of discourse*—the range of objects x relevant to the logical argument we are planning to carry through. Such freedom will commonly diminish by one the required number of terms, as we saw a few pages back in the kangaroo example. We interpret a term letter 'F', then, by settling which ones of those objects 'F' is to be construed as true of. A valid term schema, then, is one that will come out true of *all* objects of *any* chosen universe under *all* interpretations, within that universe, of its term letters.

A consistent term schema is true of *some* objects of *some* universes under *some* interpretations of its term letters. One term schema implies another when, in every universe of discourse, every interpretation makes the second schema come out true of everything that the first comes out true of.

Since Boolean term schemata have just the structure of the old truth-functional schemata, we can test them for validity, consistency, and implication by truth-value analysis as if the 'F', 'G', etc. were 'p', 'q', etc. We can even use fell swoop where appropriate. If it seems odd to substitute '$\top$' and '$\bot$' thus for term letters, then think of these tests as applying in reality not to the term schemata themselves but to the sentence schemata that would be got from the term schemata by restoring 'x' as in (1).

We shall make good use of the following law.

(i) If a Boolean term schema is consistent, then in any universe containing a given object there is an interpretation of term letters that makes the schema come out true of that object.

For, being consistent, the schema resolves to '$\top$' under some substitution of '$\top$' and '$\bot$' for the term letters. We have then merely to interpret each term letter as true or false of the given object according as its substitute was '$\top$' or '$\bot$'.

Since consistency is invalidity of the negation, (i) can also be put thus:

(ii) If a Boolean term schema is not valid, then in any universe containing a given object there is an interpretation of term letters that makes the schema come out false of that object.

Since failure of implication means consistency of the one schema with the negation of the other, (i) has also this further corollary:

(iii) If one Boolean term schema fails to imply another, then in any universe containing a given object there is an interpretation of term letters that makes the one schema true of that object and the other false of it.

We proceed now to a generalization of (i):

(iv) If $A_1, \ldots, A_n$ are Boolean term schemata and each of them separately is consistent, then in any universe containing distinct objects $x_1, \ldots, x_n$ there is an interpretation of term letters that simultaneously makes A_1 true of x_1, and A_2 true of x_2, and so on.

The reasoning is as follows. As in (i), we interpret the term letters as true or false of x_1 in such a way as to make A_1 come out true of x_1. But this leaves us free still to interpret those same term letters as true or false, as we please, of x_2; and we do so in such a way as to make A_2 come out true of x_2. We continue thus.

Example: Take A_1, A_2, and A_3 as '$F\bar{G}$', '$G\bar{H}$', and '$\bar{F}H$', and take x_1, x_2, and x_3 as Tom, Dick, and Harry. Substitution of 'T' for 'F' and '$\perp$' for 'G' makes '$F\bar{G}$' resolve to 'T'; so interpret 'F' and 'H' as true of Tom, and 'G' not. Similarly, in view of '$G\bar{H}$', interpret 'F' and 'G' as true of Dick, and 'H' not. Similarly, in view of '$\bar{F}H$', interpret 'G' and 'H' as true of Harry, and 'F' not. So 'F' is interpreted in the universe of Tom, Dick, and Harry as true of Tom and Dick; 'G' is interpreted as true of Dick and Harry; and 'H' is interpreted as true of Tom and Harry.

Having recorded for future reference these four laws of Boolean term schemata, we turn to Boolean schemata of a further kind. Let us write '$\exists$' for 'there are', in the sense 'there is at least one'. It is a prefix which, applied to a general term, produces a statement; thus 'There are black swans', schematically '$\exists FG$'. A schema consisting, like '$\exists FG$', of '$\exists$' followed by a Boolean term schema, will be called a *Boolean existence schema*.

The schematism of the example about the class of '00, at the end of the preceding chapter, goes over now into this:

(2) $\quad -\exists FG\bar{H} \rightarrow \exists F\bar{G}, \quad -\exists F\bar{G} \vee -\exists F\bar{H}, \quad -\exists FH\bar{G} \rightarrow \exists F\bar{H}G.$

Schemata such as these three, which are truth functions of Boolean existence schemata, will be called *Boolean statement schemata*.

Such schemata are counted *valid* when they come out true under all interpretations in all *nonempty* universes. It is convenient now to make

this exception of the empty universe because there are among the Boolean statement schemata, unlike the term schemata, certain ones such as '$\exists F \vee \exists \bar{F}$' which fail for the empty universe but are valid and useful otherwise. Usually the universe relative to which an argument is being carried out is already known or confidently believed not to be empty, so that the failure of a schema in the sole case of the empty universe is usually nothing against it from a practical point of view.

Usually, indeed, the universe wanted in arguments worthy of Boolean schemata is known or believed to have not merely some members, but many. In the definition of validity, then, instead of saying 'every nonempty choice of universe' why not say 'every choice of universe of more than eleven members'? The reason for not doing so is that *no* new schemata would be added to the category of valid schemata by such a liberalization of the definition. If a schema can be falsified by some interpretation of '*F*', '*G*', etc. in some small but not empty universe, it can be falsified also in bigger universes without end.

The reasoning is as follows. We are supposing a small but not empty universe; let x be an object in it. We are supposing, further, an interpretation of '*F*', '*G*', etc. in that universe that falsifies a given schema. Add, then, as many new objects to the universe as you like, and interpret '*F*', '*G*', etc. in the enlarged universe by letting all the new objects go along with x. That is, take '*F*' as true of all or none of the new objects according as '*F*' was interpreted as true or false of x; and similarly for '*G*' and further letters. Then the new objects will be indistinguishable from x so far as the interpretation of our schema is concerned; the schema will be falsified now as before.

Happily, thus, the only exception that there is any occasion to make in defining validity is that of the empty universe. But we must not underestimate this exception. Occasionally an argument may most conveniently be carried through under a choice of universe whose nonemptiness is open to question. Still there is no need to cover the empty universe in our general theory of validity, because the question whether a schema holds for the empty universe is easily handled as a separate question. To decide whether a schema comes out true for the empty universe we have merely to put '$\perp$' for the existential clauses and resolve.

Just as a Boolean statement schema counts as valid when true under all interpretations in all nonempty universes, so it counts as *consistent* when true under some interpretation in some nonempty universe. One schema implies another if, in every nonempty universe, every interpretation that makes the one true makes the other true. As before, thus, inconsistency is validity of the negation and implication is validity of the conditional.

Equivalence is, as before, mutual implication; also validity of the biconditional.

We proceed with a few laws about Boolean existence schemata.

(v) A Boolean existence schema is valid if and only if its term schema is valid.

For, if the term schema is valid, then every interpretation makes it true of everything, in any chosen universe, and therefore true of something in any nonempty universe. If on the other hand the term schema is not valid, then by (ii) there is an interpretation that makes the term schema come out false of the sole object in some universe, and the existence schema under that interpretation comes out false.

(vi) A Boolean existence schema is consistent if and only if its term schema is consistent.

For, if the term schema is consistent, then by (i) there is an interpretation that makes the term schema come out true of something, and thus makes the existence schema come out true. If on the other hand the term schema is inconsistent, every interpretation makes it false of all objects and thus falsifies the existence schema.

(vii) One Boolean existence schema implies another if and only if the one term schema implies the other.

For, suppose the one term schema implies the other. Then any interpretation that makes the one term schema true of something will make the other true of the same thing. Then any interpretation that makes the one existence schema true will make the other true; so the one existence schema implies the other. Next suppose that the one term schema does not imply the other. Then, by (iii), there is an interpretation that makes the one term schema come out true of the sole member of some universe and makes the other come out false of it. This interpretation makes the one existence schema come out true and the other false. So the one existence schema does not imply the other.

Another remarkable thing about Boolean existence schemata is that they cannot conflict with one another. The conjunction '$\exists F\bar{G} \,.\, \exists G\bar{H} \,.\, \exists \bar{F}H$' is consistent despite the conflicts among '$F\bar{G}$', '$G\bar{H}$', and '$\bar{F}H$'. We have this:

(viii) A conjunction of Boolean existence schemata is consistent so long as each of them separately is consistent.

For, when each of the existence schemata is self-consistent, the term schema in each is self-consistent by (vi). In view of (iv), then, there is in some universe an interpretation that simultaneously makes each of these term schemata come out true of one or another object, and this is an interpretation that makes the existence schemata all true together.

Important use will be made in the next chapter of this somewhat more complex law:

(ix) A Boolean existence schema is implied by a conjunction of Boolean existence schemata only if it is implied by one of them in isolation.

For, imagine Boolean term schemata in place of 'A', 'A_1', . . . , 'A_n' and suppose that no one of the existence schemata '$\exists A_1$', '$\exists A_2$', etc. separately implies '$\exists A$'; then we want to show that the conjunction '$\exists A_1 \,.\, \exists A_2 \,.\,.\,.\,.\, \exists A_n$' does not imply '$\exists A$' either. By (vii), '$A_1$' does not imply '$A$'; so '$A_1\bar{A}$' is consistent. So is '$A_2\bar{A}$', so is '$A_3\bar{A}$', etc. So, by (iv), there is in a universe of n objects $x_1, \ldots, x_n$ an interpretation of term letters that simultaneously makes '$A_1\bar{A}$' true of x_1, and '$A_2\bar{A}$' true of x_2, and so on. On this interpretation 'A' is true of none of the objects $x_1, \ldots, x_n$ in the universe, so '$\exists A_1$', '$\exists A_2$', etc. all come out true and '$\exists A$' false. Therefore '$\exists A_1 \,.\, \exists A_2 \,.\,.\,.\,.\, \exists A_n$' does not imply '$\exists A$'.

We turn finally to an obvious but important law to the effect that '$\exists$' is distributive over alternation. This is valid:

(3) $$\exists(F \vee G) \leftrightarrow .\ \exists F \vee \exists G.$$

Clearly, some persons are either saints or geniuses if and only if some are saints or some are geniuses. It is not excluded, of course, that there be both saints and geniuses, nor is it excluded that some of the saints be geniuses.

The obviousness of (3) must not be allowed to encourage the thought that '$\exists$' is distributive also over conjunction. (vii) assures us rightly enough that '$\exists FG$' implies '$\exists F \,.\, \exists G$', but the converse fails. There are round things and square, after all, but no round squares.

If this warning is unneeded, perhaps the next is more to the purpose: the combination '$- \exists$' is not distributive even over alternation. Thus take

'F' as 'horse' and 'G' as 'unicorn'. Then '$\exists G$' is false and therefore '$- \exists F \vee - \exists G$' is true; on the other hand '$- \exists(F \vee G)$' is false, since '$F \vee G$' is true of the horses.

Certainly (3) allows us to do some distribution in '$- \exists(F \vee G)$'. It gives:

$$\begin{aligned} - \exists(F \vee G) &\leftrightarrow - (\exists F \vee \exists G) \\ &\leftrightarrow . - \exists F \,.\, - \exists G. \end{aligned}$$

HISTORICAL NOTE: The idea of changeable universes goes back to DeMorgan (1846), as does the phrase 'universe of discourse'. Exclusion of the empty universe was implicit already in Frege's logic of 1879. The argument that was set forth above, for excluding the empty universe, is in Hilbert and Ackermann (1938, p. 92) and no doubt in earlier works. The sign '$\exists$' and its distribution law (3) are from Peano, 1895.

EXERCISES

1. Schematize, in the style of (2), the five examples in Exercise 1 of Chapter 17.

2. In view of (ix) and (vii), you can decide whether the schema:

$$\exists(F \leftrightarrow G) \,.\, \exists F \bar{G} \,.\, \exists(\bar{F} \vee \bar{G})$$

implies '$\exists \bar{F} G$'. How? Do so.

19
TESTS OF VALIDITY

We are now in a position to test any Boolean statement schema for validity. I shall state and justify the tests case by case, according to the form of the schema. In all cases the tests reduce to truth-value analyses of Boolean term schemata.

(i) A Boolean existence schema is valid if and only if its term schema is valid.

This was (v) of the preceding chapter.

(ii) The negation of a Boolean existence schema is valid if and only if the term schema is inconsistent.

For, the negation is valid if and only if the existence schema itself is inconsistent, and so, by (vi) of the preceding chapter, if and only if the term schema is inconsistent.

(iii) An alternation of negations of Boolean existence schemata is valid if and only if one of those negations meets the above validity test.

For, the alternation of negations amounts to the negation of the conjunction of the existence schemata; and this is valid just in case the conjunction is inconsistent. But, by (viii) of the preceding chapter, this conjunction is inconsistent just in case some one of the existence schemata is inconsistent, and hence just in case some one of the negations of existence schemata is valid.

The next test concerns the *existential conditional,* by which I mean a conditional whose antecedent is a Boolean existence schema or a conjunction of such and whose consequent is a Boolean existence schema.

(iv) An existential conditional is valid if and only if the Boolean term schema in one of the existence schemata in the antecedent implies the Boolean term schema in the consequent.

For, the existential conditional is valid if and only if the antecedent implies the consequent. By (ix) of the preceding chapter, this will happen if and only if one of the existence schemata in the antecedent implies the consequent. By (vii) of the preceding chapter, this will happen if and only if the one term schema implies the other.

The next one goes without saying.

> (v) A conjunction of Boolean statement schemata of any of the forms covered in (i)–(iv) is valid if and only if each of them comes out valid under the tests (i)–(iv).

For completeness of our test of validity of Boolean statement schemata it now remains only to show that every Boolean statement schema can be got over into one of the five forms covered in (i)–(v), by performing equivalence transformations upon its parts. This recourse to interchangeable parts depends upon a law of interchange similar to what we saw in Chapter 9. We saw there that any component of a truth-functional schema can be supplanted by any equivalent without disturbing the validity or consistency of the whole and without affecting its relations of implication and equivalence to other schemata. Such interchange works equally for Boolean statement schemata, and for much the same reason. It worked for pure truth functions because the truth value of the compound statement depends on no features of the component statements except their truth values. And it works now for truth functions and '$\exists$' because the truth value of the compound statement depends on no features of the component statements and terms except their being true or false or true of something or of nothing.

To transform any given Boolean statement schema S into one of the five forms, we begin by transforming S into conjunctional normal form as of Chapter 12. For this purpose we treat each existence schema in S as if it were a single statement letter 'p', 'q', etc. This brings S into the form of a conjunction of alternations. Each of these alternations is an alternation of existence schemata or of negations of existence schemata or both. However, there is never any need to rest with an alternation of existence schemata, thanks to the distributivity law, (3) of the preceding chapter. Transformed by this law, each of our alternations will be an alternation of (one or more) negations of existence schemata and at most one existence schema. If it is an alternation of one such negation, it falls under (ii); if of several, (iii). If it is an alternation of no negations and one affirmative existence schema, it falls under (i). If finally it is an alternation of one or

more negations of existence schemata and one existence schema, it amounts to an existential conditional and falls under (iv); for, '$\bar{p}_1 \vee \bar{p}_2 \vee \ldots \vee \bar{p}_n \vee q$' is equivalent to '$p_1 p_2 \ldots p_n \rightarrow q$'.

For a simple illustration let us apply the test to the third syllogism in Chapter 16. Rendered as Boolean statement schemata, the premises of this syllogism are '$- \exists G\bar{H}$' and '$\exists FG$' and the conclusion is '$\exists FH$'. So the inclusive Boolean statement schema to test for validity is:

$$- \exists G\bar{H} \,.\, \exists FG \,.\!\rightarrow\, \exists FH.$$

Toward converting this to conjunctional normal form, we translate the conditional into alternation and negation.

$$-(- \exists G\bar{H} \,.\, \exists FG) \vee \exists FH.$$

Conversion into conjunctional normal form takes one more step:

$$\exists G\bar{H} \vee - \exists FG \vee \exists FH.$$

Then we fuse the two affirmative existence schemata.

$$- \exists FG \vee \exists(G\bar{H} \vee FH).$$

Rendered as an existential conditional, this becomes:

$$\exists FG \rightarrow \exists(G\bar{H} \vee FH).$$

By the criterion (iv), finally, this is valid; for a fell swoop shows that 'FG' implies '$G\bar{H} \vee FH$'.

Let us apply the test to the argument regarding the class of '00. It got as far as the formulation (2) midway in the preceding chapter. To establish this implication we want to show the validity of the conditional:

$$- \exists FG\bar{H} \rightarrow \exists F\bar{G} \,.\, - \exists F\bar{G} \vee - \exists F\bar{H} \,.\!\rightarrow.\, - \exists FH\bar{G} \rightarrow \exists F\bar{H}G.$$

Appreciable labor can commonly be saved if, to begin with, we permute the letters of each term into alphabetical order. This enhances the matching of terms and, therewith, the chances of simplification. We get this:

$$- \exists FG\bar{H} \rightarrow \exists F\bar{G} \,.\, - \exists F\bar{G} \vee - \exists F\bar{H} \,.\!\rightarrow.\, - \exists F\bar{G}H \rightarrow \exists FG\bar{H}.$$

The stages of conversion into conjunctional normal form are left to the reader. If patently valid clauses of the type '$\bar{p} \vee q \vee p$' are deleted as they arise, this is what the reader is apt to end up with:

$$- \exists F\bar{G} \vee \exists F\bar{H} \vee \exists F\bar{G}H \vee \exists FG\bar{H}.$$

When we fuse the three affirmative existence schemata, this becomes:

$$- \exists F\bar{G} \vee \exists(F\bar{H} \vee F\bar{G}H \vee FG\bar{H}).$$

Rendered as an existential conditional, this becomes:

$$\exists F\bar{G} \rightarrow \exists(F\bar{H} \vee F\bar{G}H \vee FG\bar{H}).$$

By the criterion (iv), finally, this is valid; for a fell swoop shows that '$F\bar{G}$' implies '$F\bar{H} \vee F\bar{G}H \vee FG\bar{H}$'.

Now one more example:

$$- \exists F\bar{G}H \,.\, \exists F \rightarrow \exists G \,.\vee\, \exists F\bar{G}.$$

Paraphrasing the conditional into alternation and negation, we get:

$$- \exists F\bar{G}H \,.\, - \exists F \vee \exists G \,.\vee\, \exists F\bar{G},$$

which distributes to conjunctional normal form thus:

$$- \exists F\bar{G}H \vee \exists F\bar{G} \,.\, - \exists F \vee \exists G \vee \exists F\bar{G}.$$

Rendering this as a conjunction of existential conditionals, we have:

$$\exists F\bar{G}H \rightarrow \exists F\bar{G} \,.\, \exists F \rightarrow \exists(G \vee F\bar{G}).$$

Here (v) is applicable; we have to see that both conditionals are valid, if the whole is to qualify as valid. Both conditionals do, by (iv), prove valid; for we find by fell swoop that '$F\bar{G}H$' implies '$F\bar{G}$' and 'F' implies '$G \vee F\bar{G}$'.

A decision procedure for validity is of course a decision procedure at the same time for consistency, implication, and equivalence, since inconsistency, implication, and equivalence are validity of the negation, the conditional, and the biconditional.

The test of validity which is now before us always proceeds, we might say, by transforming a Boolean statement schema into a conjunction of genuine or defective existential conditionals; for the schemata covered in (i), (ii), and (iii) amount to defective existential conditionals, lacking antecedent or consequent. So the method may be called the *method of existential conditionals*.

A different validity test for Boolean statement schemata, due to Herbrand, I shall call the *cellular* method. In practice it tends to be more laborious than that of existential conditionals, but it is worth noting because it is so simply formulated. It operates with *cellular* existence schemata. A cellular existence schema in n given letters 'F', 'G', etc. has as its term schema just a conjunction of n literals in alphabetical order, each being one of those letters or its negation. Thus the cellular existence schemata in 'F' are '$\exists F$' and '$\exists \bar{F}$'; those in 'F' and 'G' are '$\exists FG$', '$\exists F\bar{G}$', '$\exists \bar{F}G$', and '$\exists \bar{F}\bar{G}$'; those in 'F', 'G', and 'H' are '$\exists FGH$', '$\exists FG\bar{H}$', and six more; and so on. Each corresponds thus to an uncut cell of the Venn diagram and says there is something in it.

Note next that any Boolean existence schema, say in n letters, can be transformed into an alternation of cellular existence schemata in those n letters. The method is as follows. Given an existence schema, put its term into developed alternational normal form (see Chapter 10) and then distribute '$\exists$' through the alternation. For instance, given '$\exists[\bar{F}(G \vee H)]$', we first put its term into alternational normal form, getting '$\exists(\bar{F}G \vee \bar{F}H)$', and then we develop it, getting:

$$\exists(\bar{F}GH \vee \bar{F}G\bar{H} \vee \bar{F}GH \vee \bar{F}\bar{G}H),$$

or, dropping the duplication,

$$\exists(\bar{F}GH \vee \bar{F}G\bar{H} \vee \bar{F}\bar{G}H).$$

Finally, distributing '$\exists$', we have:

$$\exists\bar{F}GH \vee \exists\bar{F}G\bar{H} \vee \exists\bar{F}\bar{G}H.$$

It follows that any Boolean statement schema S, say in n letters, can be rendered as a truth function of cellular existence schemata in those n letters. First, by a trivial move, each existence schema appearing in S can be made to contain all n letters. For, if an existence schema lacks 'H', we

can just append '$H \vee \bar{H}$' to its term schema by conjunction. Thereupon, by the method of the preceding paragraph, we proceed to render each existence schema as an alternation of cellular existence schemata in the n letters.

Now the validity test that was going to be so simply formulable is just this:

> (vi) A truth function of cellular existence schemata in n given letters is a valid Boolean statement schema if and only if it resolves to '$\top$' under all assignments of '$\top$' and '$\bot$' to those cellular schemata (disregarding the assignment of '$\bot$' to all 2^n cellular schemata).

I shall not pause for a full proof of (vi), for it can quickly be made plausible by the following reflection upon Venn diagrams. When we undertake to interpret a Boolean statement schema, the only information that is going to matter to its truth or falsity is information as to which cells of the diagram are occupied and which are empty. Moreover, interpretation of any one cell as occupied or empty is independent of interpretation of other cells as occupied or empty. So the only thing that can make for validity of a Boolean statement schema, once the schema is resolved into the cellular existence schemata, is the external truth-functional structure by which it is built of those cellular existence schemata. (Exception: since we chose to define validity without responsibility for the empty universe, the parenthetical exemption is appended to (vi).)

Example:

$$\exists FG \vee \exists FH \mathrel{.\!\!\rightarrow} \exists[F(G \vee H)].$$

First we supply missing letters to all its existence schemata.

$$\exists[FG(H \vee \bar{H})] \vee \exists[FH(G \vee \bar{G})] \mathrel{.\!\!\rightarrow} \exists[F(G \vee H].$$

Then we expand the terms into normal form, holding to alphabetical order.

$$\exists(FGH \vee FG\bar{H}) \vee \exists(FGH \vee F\bar{G}H) \mathrel{.\!\!\rightarrow} \exists(FG \vee FH).$$

Then we develop the terms, holding to alphabetical order and dropping a duplicate.

$$\exists(FGH \vee FG\bar{H}) \vee \exists(FGH \vee F\bar{G}H) .\!\rightarrow \exists(FGH \vee FG\bar{H} \vee F\bar{G}H).$$

Then we distribute '$\exists$'.

$$\exists FGH \vee \exists FG\bar{H} \vee \exists FGH \vee \exists F\bar{G}H .\!\rightarrow\!. \exists FGH \vee \exists FG\bar{H} \vee \exists F\bar{G}H.$$

The truth-functional validity is visible.

The stipulation of alphabetical order in the definition of cellular existence schema is no mere labor saver now; it is essential to the soundness of (vi). If in this example we had written '$\exists FGH$' at some points and '$\exists FHG$' at others, the truth-functional validity of the outward structure would have been lost; the parts '$\exists FGH$' and '$\exists FHG$' would have fared like independent sentence letters 'p' and 'q'.

HISTORICAL NOTE: The first decision procedure for this much of logic was Löwenheim's (1915), but the first reasonably manageable one was Behmann's (1922). The method of existential conditionals is somewhat reminiscent of Behmann's. The cellular method stems rather from Herbrand (1930).

EXERCISES

1. Test each of the following schemata for validity by the method of existential conditionals.

$$\exists JF \rightarrow \exists(F\bar{G} \vee JG \vee GH),$$
$$\exists GJ \,.\, \exists GK .\!\rightarrow \exists(FGH \vee \bar{F}J \vee \bar{H}K).$$

2. Schematize the following statement and test for consistency (by testing the negation for validity) according to the method of existential conditionals.

> Some who take logic and Latin take neither physics nor Greek, but all who take either Latin or chemistry take both logic and Greek.

3. Five inferences were schematized in Exercise 1 of Chapter 18. Check

them all, by testing the appropriate conditionals for validity according to the method of existential conditionals.

4. Check this inference by the same method.

   Nobody attended but parents of the cast.
   If all parents of the cast were bored, then no parents of the cast have a taste for high-class entertainment.
   ∴ If any attended who have a taste for high-class entertainment, then some parents of the cast were not bored.

5. Repeat all four exercises using the cellular method.

20
SOME BOOLEAN INCIDENTALS

Boolean term schemata are truth functions of term letters—thus '$\bar{F}$', 'FG', '$F \vee G$', '$F \rightarrow G$', '$F \leftrightarrow G$', and upward by iteration. Boolean existence schemata are term schemata with '$\exists$' prefixed, and Boolean statement schemata are these and the truth functions of them. This meager inventory is already redundant, as we know; '$-(\bar{F}\bar{G})$' would serve for '$F \vee G$', and '$-(F\bar{G})$' for '$F \rightarrow G$', and '$-(F\bar{G}) \; -(\bar{F}G)$' for '$F \leftrightarrow G$'. Conversely, convenience can be served by further redundancy. Along with '$\exists F$' for 'there are F' we may conveniently adopt '$\forall F$' for 'all is F', 'there are nothing but F'. Convenient further notations are that of *inclusion*, '$F \subseteq G$', 'all F are G'; that of *coextensiveness*, '$F \equiv G$', 'all and only F are G'; and that of *proper inclusion*, '$F \subset G$', 'all F are G but not vice versa'. All these are easily reducible to the previous notations, thus:

$$\forall F \leftrightarrow - \exists \bar{F},$$

$$F \subseteq G \mathrel{.\leftrightarrow} \forall(F \rightarrow G) \leftrightarrow - \exists(F\bar{G}).$$

$$F \equiv G \mathrel{.\leftrightarrow} \forall(F \leftrightarrow G) \leftrightarrow .F \subseteq G \,.\, G \subseteq F \leftrightarrow .- \exists(F\bar{G}) \,.\, - \exists(\bar{F}G).$$

$$F \subset G \mathrel{.\leftrightarrow.} F \subseteq G \,.\, -(G \subseteq F) \leftrightarrow .- \exists(F\bar{G}) \,.\, \exists(\bar{F}G).$$

In (3) of Chapter 18 we noted that '$\exists$' is distributive through alternation. Similarly '$\forall$' is distributive through conjunction:

$$(1) \qquad \forall FG \leftrightarrow . \forall F \,.\, \forall G.$$

If the validity of (1) were less obvious, one might check it by translating it into the previous notation, as just explained, and then applying the method of existential conditionals. Similarly for the validity of this conditional:

$$(2) \qquad \forall F \rightarrow \exists F.$$

Universality, ubiquity, implies existence.

It is natural and convenient to record equivalences and implications between statement schemata by setting down valid biconditionals and conditionals as above. These do not themselves affirm equivalence or implication, but their validity constitutes equivalence and implication. Now the signs '$\equiv$' and '$\subseteq$' of coextensiveness and inclusion perform that same service for term schemata. That 'FG' is equivalent to '$-(\bar{F} \vee \bar{G})$', and implies '$F \vee G$' is reflected in the validity of the corresponding statement schemata of coextensiveness and inclusion:

$$FG \equiv -(\bar{F} \vee \bar{G}), \qquad FG \subseteq F \vee G.$$

The word *functor,* grammatical in import but logical in habitat, becomes useful at this point. A functor is a sign that attaches to one or more expressions of given grammatical kind or kinds to produce an expression of a given grammatical kind. The negation sign is a functor that attaches to a statement to produce a statement and to a term to produce a term. The alternation sign, conditional sign, and biconditional sign are functors that join statements in pairs to produce statements and join terms in pairs to produce terms. '$\exists$' is a functor that attaches to a term to form a statement; so is '$\forall$'. Finally '$\subseteq$', '$\subset$', and '$\equiv$' are functors that join terms to form statements. Functors of this last kind are called *copulas;* in particular '$\subseteq$' means 'are'.

Clearly then '$F \subseteq G$' and '$F \equiv G$' do not correspond to '$p \rightarrow q$' and '$p \leftrightarrow q$' in the way that 'FG' and '$F \vee G$' correspond to 'pq' and '$p \vee q$'. What correspond to '$p \rightarrow q$' and '$p \leftrightarrow q$' in the way that the term schemata 'FG' and '$F \vee G$' correspond to 'pq' and '$p \vee q$' are rather the term schemata '$F \rightarrow G$' and '$F \leftrightarrow G$'. The sentence schemata '$F \subseteq G$' and '$F \equiv G$' are not to be confused with the term schemata '$F \rightarrow G$' and

'$F \leftrightarrow G$'; they are related to them rather, we saw, as '$\forall(F \rightarrow G)$' and '$\forall(F \leftrightarrow G)$'.

The Boolean logic of terms that is now before us is substantially what has traditionally been conceived of as an algebra of classes and called Boolean algebra. The letters 'F', 'G', etc., used just now as schematic letters standing for general terms, would from that point of view be counted rather as standing for class names, or as variables ranging over classes. The copula '$\equiv$' gives way on that approach to simple identity, '$=$'; for classes are identical when they have the same members. Traditionally indeed equations predominate in Boolean algebra; an algebraic look prevails. Thus instead of '$\exists F$' and '$\forall F$' we get '$F \neq \Lambda$' and '$F = \mathrm{V}$', Λ being the empty class and V the universe. The intent is no longer logic proper, but the substantive theory of abstract objects of a special sort: classes. This dissociation from truth-function logic is accentuated by further departures in notation; thus '$F \rightarrow G$' and '$F \leftrightarrow G$' are not written, and instead of the alternation '$F \vee G$' and conjunction 'FG' of terms we get '$F \cup G$' and '$F \cap G$', said to designate the *union* and *intersection* of classes.

Waiving this last departure, we may describe a *Boolean equation* as an equation each side of which is one of the letters 'Λ', 'V', 'F', 'G', etc., or is compounded of them by negation, alternation, conjunction.

In the ordinary algebra of numbers, it will be recalled, any equation can be converted to a form in which one side is '0'. Now it happens likewise that in Boolean algebra any equation can be converted to a form in which one side is 'Λ'. For, '$F = G$' means that the classes F and G have the same members; hence that there is nothing in $F\bar{G}$ nor in $\bar{F}G$; and hence that $F\bar{G} \vee \bar{F}G = \Lambda$.

Equally, if we prefer, we can always make one side 'V'; for '$F = G$' is equivalent also to '$FG \vee \bar{F}\bar{G} = \mathrm{V}$'.

The formulas of Boolean algebra are the Boolean equations and the truth functions of them. We can test any such formula for validity by translating it into a Boolean schema as of Chapter 18 and testing it as in Chapter 19. So let us see how to translate. We saw how to transform any equation into the form '$\Lambda = \ldots$'; and then '$\Lambda =$' can be translated as '$- \exists$'. If there are residual occurrences of 'Λ' or 'V', they can be resolved out truth-functionally in the manner of '$\bot$' and '$\top$' or, more clumsily, rewritten as '$F\bar{F}$' and '$F \vee \bar{F}$'. Any truth function of Boolean equations goes over thus into a truth function of Boolean existence schemata, ready for the routine of Chapter 19.

Is there more than a notational difference between a Boolean logic of

terms, as in Chapters 18–19, and this algebra of classes? Nominally there is the difference between '*F*' as depicting a general term and '*F*' as depicting a class name. There is the difference between the general term 'man', or 'is a man', and the class name 'mankind'. (See Chapter 14.) The general term is true of each of various individuals, men. The class name is a name of one abstract object, the class of men. By means of the general term we may speak generally of men and not raise the philosophical question whether, besides those various men, there is an additional object which is the class of them. The class name does raise that question, and it solicits an affirmative answer.

The question is not an empty one. In Chapters 46 and 47 we shall see that the assumption of classes can add materially to what can be formulated in a theory, even when what we are formulating is not itself about classes. Where we are interested in economy of means, and of keeping account of the assumptions needed in proving or defining something, we must take note of whether classes figure in the apparatus; sometimes they do and sometimes not.

In the case of Boolean algebra, still, we are free to disavow the assumption of classes, representing it as a mere manner of speaking; for we have seen how to translate Boolean algebra into the more innocent idiom of a Boolean logic of general terms. When the assumption of classes materially contributes to a theory, and cannot be thus facilely disavowed, the classes are assumed among the very objects that the general terms themselves are true of. They are assumed as values of the variables—to anticipate a notion which we shall take up in the next chapter. Meanwhile we may best look upon the Boolean algebra of classes as only a simulated little theory of classes and a figurative rendering of what is really nothing more than the Boolean logic of general terms.

Classes are also called sets. A speculative branch of mathematics called set theory stems from Cantor (fl. 1890). It is occupied largely with problems of infinity, and it treats of classes in substantial ways that are by no means to be dismissed like the Boolean class algebra as mere manners of speaking. We shall glimpse a little of it in Chapters 46–48. In that rarefied air the term 'set' tends to be preferred to 'class', except in a certain technical context (see last pages of book) where the double terminology has been put to use to mark a special distinction.

In the current "new math" of the elementary schools, the exalted name of set theory is confusingly applied to the other end of the scale: to the Boolean algebra of classes, hence really the simple logic of general terms.

HISTORICAL NOTE: The inclusion signs '$\subset$' and '$\subseteq$', now current in set theory, are derived from Gergonne's use in 1816 of 'C' for containment. The Boolean algebra of classes was founded by Boole in 1847 and improved by DeMorgan and Jevons in 1860 and 1864. It was anticipated in some measure by Leibniz in the seventeenth century. Between this logic and the logic of truth functions, whose antecedents are more ancient (see Chapter 2), there are of course conspicuous parallels; and they were noted already by Leibniz and Boole. But the full relationship between the logic of truth functions and the Boolean algebra of classes was not as clear as it later became, e.g. with Behmann (1927). The relationship is best reflected by the logic of terms, Chapters 18 and 19 above, putting classes aside.

EXERCISES

1. Translate the following formulas into Boolean schemata in the sense of Chapters 18 and 19 and test them for validity by the method of existential conditionals.

$$F = F \vee G \vee H \mathrel{.\!\rightarrow\!.} F = G \vee H,$$
$$F \vee \bar{G} = G \mathrel{.\!\rightarrow\!.} F = G,$$
$$F = F \vee G \mathrel{.\!\leftrightarrow\!.} FG = G,$$
$$FG = F \vee G \mathrel{.\!\leftrightarrow\!.} F = G.$$

2. Reformulate the method of existential conditionals as a decision procedure applicable directly to truth functions of Boolean equations. (A serious exercise.)

21
THE BOUND VARIABLE

A general term may, as remarked in Chapter 14, be thought of indifferently as singular or plural and as substantive, adjective, or verb. This latitude is all very well, but we could do with more. We would like to be able to handle any statement about an object as predicating some

general term of that object, even when the sentence is complex and shows no neatly segregated term to the purpose. Already in the little exercise of Chapter 14 we saw that verbal contortions were needed to segregate the appropriate terms. In many examples, e.g.:

(1) Tom used to work for the man who murdered the second husband of Tom's youngest sister,

it can be rather a challenge to find a term which, predicated of Tom, has the same effect. We come out with:

(2) former employee of own youngest sister's second husband's murderer.

Ordinary language already affords a general expedient for such cases: the *relative clause*. The purpose of (2) is served, very nearly, by this:

(3) who used to work for the man who murdered the second husband of his youngest sister.

The new trouble is the cross-reference of 'his'; proximity steers it to the wrong 'who'. Problems of cross-reference involve us in vexatious recasting, as we so well know. There is an inelegant quasi-mathematical usage, however, that cuts through all that. It is the 'such that' construction, with a supporting variable supplanting the pronoun:

(4) x such that x used to work for the man who murdered the second husband of x's youngest sister,

or, to be prolix,

(5) x such that x used to work for the man y such that y murdered the second husband of x's youngest sister.

The 'x' and 'y' here are *bound variables:* mere devices for pronominal cross-reference. Multiple choice of letters is wanted for avoiding ambiguity of cross-reference when, as in (5), relative clauses are nested.

In the relative clause, then, rendered with 'who', 'which', 'that', or 'x such that', we have a general term that segregates what the original sentence said about the object in question. Predicating it of the object is equivalent to the original sentence. 'Tom is (someone) who enjoys music'

amounts to 'Tom enjoys music'; 'Tom is (someone) x such that x used to work etc.' amounts to (1). The insertion of 'someone' here is just a concession to grammar, adjusting an adjectival clause to substantival position.

In forming a relative or 'such that' clause, we segregate and we abstract. In (2)–(5) what (1) says of Tom is segregated from the name 'Tom' and abstracted from the sentence (1). What we thus abstract is a general term; there is no thought here of an abstract object. The clause thus abstracted may, like any general term, be true of any number of objects of whatever sorts. Abstract objects are as may be; we shall come to direct grips with them only briefly toward the end of the book.

The construction 'x such that . . . x . . .' will hereafter be rendered '$\{x$: . . . x . . .$\}$' and called *term abstraction*. The complex general term thus formed will be called an *abstract*. Just as predicating (4) of Tom simply reaffirms (1), so in general

(6) $$\{x\colon \ldots x \ldots\} y \leftrightarrow \ldots y \ldots.$$

Such reduction of a predication '$\{x$: . . . x . . .$\}y$' to '. . . y . . .' I call *concretion*. When the sentence that I have represented as '. . . x . . .' is just a simple predication, say 'x is wise', the abstract becomes merely a redundant rendering of 'wise' itself;

$$\{x\colon x \text{ is wise}\} \equiv \text{wise}.$$

Thus

(7) $$\{x\colon Fx\} \equiv F, \qquad \{x\colon Fx\} y \leftrightarrow Fy.$$

The 'x' that appeared so briefly in (1) of Chapter 18 has now come back to stay. It is the bound variable of term abstraction; but let us be more precise. It is bound *in* the abstract '$\{x$: x is wise$\}$', bound *by* the prefix 'x:', but *free* in the component expression 'x is wise' taken by itself. This expression is called an *open sentence*. It differs from a *closed sentence*, or a statement, in containing a variable in place of a name. It is neither true nor false. The analogue of a free variable in ordinary language is a pronoun for which no grammatical antecedent is expressed or understood, and the analogue of an open sentence is a clause containing such a dangling pronoun.

With abstraction at hand, the relation of the truth functions of terms to those of sentences can be rendered vividly anew:

$$\bar{F} \equiv \{x\colon Fx\}, \tag{8}$$
$$FG \equiv \{x\colon Fx \,.\, Gx\}, \tag{9}$$
$$F \vee G \equiv \{x\colon Fx \vee Gx\}, \tag{10}$$
$$F \rightarrow G \equiv \{x\colon Fx \rightarrow Gx\}, \tag{11}$$
$$F \leftrightarrow G \equiv \{x\colon Fx \leftrightarrow Gx\}. \tag{12}$$

Instead of expressing predication by juxtaposition, as in 'Fx', I shall sometimes use an epsilon as copula—particularly in connection with abstracts. Thus (6) becomes:

$$y \in \{x\colon \ldots x \ldots\} \,.\leftrightarrow \ldots y \ldots. \tag{13}$$

Whereas the copulas '$\subseteq$', '$\subset$', and '$\equiv$' join general terms, the copula '$\in$' joins a singular term to a general term; for the variable behaves grammatically like a singular term. The copulas '$\in$' and '$\subseteq$' answer to 'is' and 'are'.

The abstract and its epsilon are associated with set theory. The class of all objects x such that $\ldots x \ldots$ is called $\{x\colon \ldots x \ldots\}$ in set theory, and membership is expressed by epsilon. My adoption of these notations for the ontologically innocent relative clause and copula is of a piece with my ontologically innocent use of the Boolean formalism that has been gratuitously saddled with the burden of classes down the years.

Some readers will feel that I would do better here to render term abstraction by a distinctive sign '$\ni$' for 'such that', as I and others have sometimes done, and leave '$\{x\colon \ldots x \ldots\}$' and '$\in$' to unregenerate set theory. I now find it more instructive philosophically and more elegant logically to use the same notation for both. When the time comes to broach set theory and recognize classes as objects (Chapters 46–48), general terms will be enlisted for double duty as names of classes. Term abstraction will then resume intermittently its accustomed office of class abstraction.

HISTORICAL NOTE: Being the initial letter of the Greek copula, epsilon was adopted for the singular copula by Peano in 1889. He inverted it for the 'such that' of term abstraction because of how they cancel

each other; see (13). Peano's 'ϵ' is in universal use today for class membership, but his '$\ni$' is used nowadays rather for the noncommittal 'such that' of term abstraction, usually informally. Peano used 'ϵ' for both the copula and membership, and '$\ni$' for both term abstraction and class abstraction, much as I am venturing to do—though his was a case of genuine confusion. A notation using circumflex accents and serving unequivocally for class abstraction was used by Russell in 1908 and persisted through *Principia Mathematica* and much subsequent literature, including early editions of this book. The notation of curly braces that I am now using is usual today in set theory.

In formulating set theory in 1940 Gödel appealed informally to fictitious supplementary objects that he called *notions*. They were like classes but were not values of variables. They were just a manner of speaking, avoidable by circumlocution. In my Brazil lectures of 1942 I pressed this fiction. I presented what I called the virtual theory of classes and relations, in which I used the notation of set theory as far as I could without assuming classes. See *O Sentido da nova lógica,* §51. Martin was urging the idea concurrently. In *Set Theory and Its Logic* (1963, 1967) I made extensive use of the virtual theory of classes as an auxiliary to the real theory of classes. It proved to be a valuable auxiliary, but I persisted in presenting it as a mere prolegomenon to set theory and as a sly partial simulation thereof, failing, as we all did, to appreciate that it should stand squarely in elementary logic as a regimentation of a fundamental feature of language, the relative clause.

22
QUANTIFICATION

To say that all men are mortal, and that some books are boring, we can now write:

(1) $\forall\{x\colon \text{man } x \rightarrow \text{mortal } x\}$,

(2) $\exists\{x\colon \text{book } x \mathbin{.} \text{boring } x\}$.

Application of the functors '$\forall$' and '$\exists$' to term abstracts in this fashion proves strategic. Logic can be pursued without using the abstracts otherwise than after '$\forall$' and '$\exists$'. The converse is also true: '$\forall$' and '$\exists$' are never needed otherwise than in application to term abstracts, since, by (7) of the preceding chapter, '$\forall F$' and '$\exists F$' can always be rendered '$\forall\{x: Fx\}$' and '$\exists\{x: Fx\}$'. In short, the combinations '$\forall\{x:$' and '$\exists\{x:$', also '$\forall\{y:$', '$\exists\{y:$', etc., can be treated hereafter as simple prefixes. I shall do so, and shorten them to '$\forall x$', '$\exists x$', '$\forall y$', '$\exists y$', etc. (1) and (2) become *quantifications:*

(3) $\forall x(\text{man } x \rightarrow \text{mortal } x)$,

(4) $\exists x(\text{book } x \,.\, \text{boring } x)$.

The prefixes '$\forall x$', '$\exists x$', '$\forall y$', etc., called *quantifiers,* are the focal point of neoclassical logic.

Relations noted in Chapter 20 between '$\forall$' and '$\exists$' reappear now between quantifications:

$$\forall x\, {-}Fx \leftrightarrow {-}\,\exists x\, Fx, \qquad \exists x\, {-}Fx \leftrightarrow {-}\,\forall x\, Fx, \qquad \forall x\, Fx \rightarrow \exists x\, Fx.$$

When we switch to quantificational style, we drop the Boolean compounds '$\bar{F}$', 'FG', '$F \vee G$', '$F \rightarrow G$', and '$F \leftrightarrow G$'; they go over into term abstracts as in (8)–(12) of the preceding chapter and disappear into quantifications. Term letters thus cease to occur otherwise than with a variable appended: 'Fx', 'Gx', 'Fy'. Thenceforward we can think of these combinations as standing for sentences as unanalyzed wholes. 'Fx' stands for any open sentence containing free 'x' buried anywhere within it, perhaps in several places, as 'Tom' was in (1) of Chapter 21, and we need have no thought of segregating a term represented by 'F' itself. If we do have the thought, we can always reactivate term abstraction; the term is (4) of Chapter 21.

As of now the move to quantification is no improvement. On the contrary, Boolean schemata as of Chapters 18 and 19 are simpler and more perspicuous, unencumbered by either term abstracts or quantifiers. We have our test of validity by existential conditionals, and little is left to be desired—until we move from absolute or *monadic* terms like 'book' to relative or *dyadic* terms like 'uncle'. This is the move that complicates logic and makes for its stature as a serious subject. It is only in preparation for it that we now switch to quantifiers. But we shall have five

chapters on them before we tackle the dyadic terms for which they are ultimately wanted.

The quantificational renderings of the categorical forms **A** and **I** of Chapter 14 are illustrated in (3) and (4) above; and **E** and **O** are equally evident. To sum up:

A:	All F are G $\forall x(Fx \rightarrow Gx)$	**E**:	No F are G $\forall x(Fx \rightarrow -Gx)$
I:	Some F are G $\exists x(Fx \,.\, Gx)$	**O**:	Some F are not G $\exists x(Fx \,.\, -Gx)$

To discourage the erroneous reading '$\exists x(Fx \rightarrow Gx)$' of **I** once and for all, let us stop to see what '$\exists x(Fx \rightarrow Gx)$' really says. '$Fx \rightarrow Gx$' is equivalent to '$-Fx \vee Gx$', so '$\exists x(Fx \rightarrow Gx)$' says only that at least one object is non-F or G; and this is bound to be true, regardless of how 'F' and 'G' are interpreted, except in the one extreme case where 'F' is true of everything in the universe and 'G' is true of nothing. The form '$\exists x(Fx \rightarrow Gx)$' is so rarely false as to be seldom worth affirming.

An odd feature of ordinary language is the grammar of the words 'everything' and 'something', which behave like names except at crucial points. The statements:

Maud is a book and is boring,
Maud is a book and *Maud* is boring

are clearly interchangeable, as are:

Maud is in prose or in verse,
Maud is in prose or *Maud* is in verse;

but the statements:

(5) Something is square and round,
(6) Something is square and something is round

have opposite truth values, as do these:

(7) Everything is visible or invisible,

(8) Everything is visible or everything is invisible.

Such is the importance of *scope* in quantification. In '$\exists x(Fx \,.\, Gx)$' and '$\forall x(Fx \vee Gx)$' the scope of the initial quantifier extends to the end; in '$\exists x\, Fx \,.\, \exists x\, Gx$' and '$\forall x\, Fx \vee \forall x\, Gx$' it stops halfway. In these cases it matters. In the opposite cases it is indeed immaterial: existential quantification is distributive through alternation and universal quantification is distributive through conjunction.

(9) $\exists x(Fx \vee Gx) \leftrightarrow .\ \exists x\, Fx \vee \exists x\, Gx,$

(10) $\forall x(Fx \,.\, Gx) \leftrightarrow .\ \forall x\, Fx \,.\, \forall x\, Gx.$

To say this is just to reiterate (3) of Chapter 18 and (1) of Chapter 20 in the light of (9) and (10) of Chapter 21.

Statements of ordinary language which at first glance seem to be conjunctions or conditionals often demand interpretation as quantifications of conjunctions or conditionals. Examples are:

(11) Sadie stole something at the Emporium and exchanged it for a blouse,

(12) If Sadie wants anything she manages to get it.

These must be interpreted as quantifications:

(13) $\exists x$(Sadie stole x at the Emporium . Sadie exchanged x for a blouse),

(14) $\forall x$(Sadie wants $x \rightarrow$ Sadie manages to get x),

rather than as a conjunction and conditional:

(15) $\exists x$(Sadie stole x at the Emporium) . Sadie exchanged it for a blouse,

(16) $\exists x$(Sadie wants x) $\rightarrow$ Sadie manages to get it.

For, the 'it' of (11) clearly refers back across 'and' to 'something', and correspondingly the 'it' of (12) refers back to 'anything'. The quantifiers must be made to cover the whole compound as in (13) and (14), rather than just the first clause as in (15) and (16), so as to reach out to a lagging recurrence of 'x' in the position of 'it'.

If we think of the universe as limited to a finite set of objects $a, b, \ldots, h$, we can expand existential quantifications into alternations and universal quantifications into conjunctions; '$\exists x\, Fx$' and '$\forall x\, Fx$' become respectively:

$$Fa \vee Fb \vee \ldots \vee Fh, \qquad Fa \,.\, Fb \,.\ldots.\, Fh.$$

The distinction between (5) and (6) then comes out quite clearly; '$\exists x(Fx \,.\, Gx)$' becomes:

$$Fa \,.\, Ga \,.\vee.\, Fb \,.\, Gb \,.\vee.\ldots.\vee.\, Fh \,.\, Gh,$$

whereas '$\exists x\, Fx \,.\, \exists x\, Gx$' becomes:

$$Fa \vee Fb \vee \ldots \vee Fh \,.\, Ga \vee Gb \vee \ldots \vee Gh.$$

The distinction between (7) and (8) comes out equally clearly. Furthermore the interchangeability of $-\forall x$' with '$\exists x-$' and of '$-\exists x$' with '$\forall x-$' turns out to be a mere application of DeMorgan's laws (Chapter 10); for, '$-\forall x\, Fx$' and '$\exists x\ -Fx$' become respectively:

$$-(Fa \,.\, Fb \,.\ldots.\, Fh), \qquad -Fa \vee -Fb \vee \ldots \vee -Fh,$$

and '$-\exists x\, Fx$' and '$\forall x\ -Fx$' become respectively:

$$-(Fa \vee Fb \vee \ldots \vee Fh), \qquad -Fa \,.\, -Fb \,.\ldots.\, -Fh.$$

It thus appears that quantification could be dispensed with altogether in favor of truth functions if we were willing to agree for all purposes on a fixed and finite and listed universe $a, b, \ldots, h$. However, we are unwilling; it is convenient to allow for variations in the choice of universe. This is convenient not only because philosophers disagree regarding the limits of reality, but also because—as already noted—some logical arguments can be simplified by deliberately limiting the universe of discourse to animals or to persons or to the employees of a given firm for the space of the problem in hand. For most problems, moreover, the relevant universe comprises objects which we are in no position to list in the manner of $a, b, \ldots, h$. For many problems the universe even comprises infinitely many objects; e.g., the integers. Thus it is that quantification is here to stay.

HISTORICAL NOTE: Quantification, in a different but structurally equivalent notation, was introduced by Frege in 1879. The quantifiers were rendered as '(x)' and '$(\exists x)$' by Russell in 1908, and this notation came down through *Principia Mathematica* and subsequent literature into the previous editions of this book. Meanwhile the more symmetrical pair '$\forall x$' and '$\exists x$' has gained ground. An advantage is that the '$\forall$' can be used separately along with '$\exists$', as in Chapter 20. Some writers, wishing to stress the analogy of the two quantifiers to conjunction and alternation, have used enlarged variants '$\bigwedge$' and '$\bigvee$' of their conjunction and alternation signs instead of '$\forall$' and '$\exists$'.

EXERCISES

1. Rewrite these with help of quantifiers:

   John cannot outrun any man on the team,
   John cannot outrun every man on the team.

2. Rewrite the premises and conclusion about the slovenly and the poor (Chapter 17) with help of quantifiers. Similarly for the other examples in Chapter 17 and in Exercise 1 of Chapter 17 and Exercises 2 and 4 of Chapter 19.

3. Supposing the universe to comprise just $a, b, \ldots, h$, express these truth-functionally:

   $\exists x(Fx \vee Gx)$, $\forall x(Fx \vee Gx)$, $\forall x(Fx \,.\, Gx)$,
   $\exists x\, Fx \vee \exists x\, Gx$, $\forall x\, Fx \vee \forall x\, Gx$, $\forall x\, Fx \,.\, \forall x\, Gx$.

   Which come out equivalent to one another?

23
RULES OF PASSAGE. MONADIC SCHEMATA

In the schemata in which we have illustrated quantification thus far, the scope of each occurrence of a quantifier has been a truth function of 'Fx', 'Gx', etc. But we can be more liberal, allowing some components of the scope to be devoid of 'x' and represented simply as 'p' or 'q'.

It makes good sense, e.g., to write '$\exists x(p \,.\, Fx)$'. This should come out true under just those interpretations of 'p' and 'F' that make '$p \,.\, Fx$' true of at least one thing x. Those, of course, are just the interpretations that make 'p' true and 'F' true of something; so the two schemata:

(1) $$\exists x(p \,.\, Fx), \qquad p \,.\, \exists x\, Fx$$

are equivalent. By similar reasoning, so are the two schemata:

(2) $$\forall x(p \,.\, Fx), \qquad p \,.\, \forall x\, Fx.$$

The same may be expected of the pair:

(3) $$\exists x(p \vee Fx), \qquad p \vee \exists x\, Fx$$

and again of the pair:

(4) $$\forall x(p \vee Fx), \qquad p \vee \forall x\, Fx.$$

These expectations can be confirmed by direct reasoning rather like what led to (1) and (2). But also, and perhaps more instructively, we can derive these equivalences from the others. Take (3). '$\exists x(p \vee Fx)$' amounts to '$-\forall x\, -(p \vee Fx)$' and hence, by DeMorgan's law, to '$-\forall x(\bar{p} \,.\, -Fx$'. But the '$\forall x(\bar{p} \,.\, -Fx)$' in this is equivalent by (2) to '$\bar{p} \,.\, \forall x\, -Fx$'; so the whole '$-\forall x(\bar{p} \,.\, -Fx)$' becomes '$-(\bar{p} \,.\, \forall x\, -Fx)$', or, by DeMor-

gan's law, '$p \vee - \forall x\ -Fx$', which is to say, '$p \vee \exists x\ Fx$'. Precisely parallel steps carry us from (1) to (4).

The equivalences:

(5) $\exists x(p \rightarrow Fx), \quad p \rightarrow \exists x\ Fx,$

(6) $\forall x(p \rightarrow Fx), \quad p \rightarrow \forall x\ Fx$

proceed immediately from (3) and (4), respectively, since '$p \rightarrow$' amounts to '$\bar{p} \vee$'. But the two further equivalences that the reader is now perhaps on the point of expecting are not forthcoming. What are right are rather the following two; note the curious twist.

(7) $\exists x(Fx \rightarrow p), \quad \forall x\ Fx \rightarrow p.$

(8) $\forall x(Fx \rightarrow p), \quad \exists x\ Fx \rightarrow p.$

For all its oddity, (7) is easily got from (3). '$\exists x(Fx \rightarrow p)$' amounts to '$\exists x(-Fx \vee p)$', hence to '$\exists x\ -Fx \vee p$', hence to '$- \forall x\ Fx \vee p$', hence to '$\forall x\ Fx \rightarrow p$'. (8) proceeds similarly from (4).

In future applications of (1)–(4) a habit already acquired in Part I should be kept: that of ignoring order in conjunctions and in alternations. Thus, in the argument above I equated '$\exists x(-Fx \vee p)$' to '$\exists x\ -Fx \vee p$' on the strength of (3), ignoring the fact that (3) shows 'p' rather to the left of '$\vee$' and 'Fx' to the right.

Quantifiers are subject, in point of grouping, to the same convention which governs the negation sign (Chapter 4): a quantifier applies to the shortest possible ensuing sentence or schema. In '$\exists x(Fx \rightarrow p)$' the quantifier is required by the parentheses to apply to the whole conditional; in '$\exists x\ Fx \rightarrow p$', on the other hand, the quantifier is understood as lying within the antecedent and applying only to it. The distinction between '$\exists x(Fx \rightarrow p)$' and '$\exists x\ Fx \rightarrow p$' can be reproduced in words thus:

There is something x such that if Fx then p,
If there is something x such that Fx then p.

Similarly for '$\forall x(Fx \rightarrow p)$' and '$\forall x\ Fx \rightarrow p$':

Everything x is such that if Fx then p,
If everything x is such that Fx then p.

To appreciate the difference of meaning in this latter pair, take 'Fx' as

'x contributes' and 'p' as 'I'll be surprised'. Then '$\forall x(Fx \rightarrow p)$' says of each person, even the most generous, that if he contributes I'll be surprised. It says that I foresee no contribution at all. On the other hand '$\forall x\ Fx \rightarrow p$' reflects no such cynicism; it says only that I'll be surprised (as who wouldn't?) in the extraordinary event that they all contribute. In ordinary language the difference is drawn by opposing 'any' to 'every'; '$\forall x(Fx \rightarrow p)$' says that if anyone contributes I'll be surprised, while '$\forall x\ Fx \rightarrow p$' says only that if everyone contributes I'll be surprised.

In general the difference in English usage between 'any' and 'every' may have struck many of us as unsystematic and even mysterious. Scope of the universal quantifier is the key to it. Where distinction is needed between broader and narrower scope, as between '$\forall x(Fx \rightarrow p)$' and '$\forall x\ Fx \rightarrow p$', the English speaker's unconscious understanding is that 'any' calls for the broader scope and 'every' for the narrower. This rule works not only in connection with the conditional but also elsewhere, notably in connection with negation. Thus take the universe of discourse as consisting of all poems. 'I do not know any poem', then, and 'I do not know every poem', call respectively for the broader and the narrower scope:

$$\forall x\ -(\text{I know } x), \qquad -\forall x(\text{I know } x).$$

The equivalences (1)–(8) are *rules of passage*. They show how to move a quantifier across a conjunction sign, an alternation sign, or a conditional sign. Rules of passage for moving quantifiers across a negation sign were noted in the preceding chapter, in these two equivalences:

(9) $$\forall x\ -Fx, \qquad -\exists x\ Fx.$$

(10) $$\exists x\ -Fx, \qquad -\forall x\ Fx.$$

The 'p' in the rules of passage schematizes a sentence devoid of 'x', such as 'I'll be surprised'. But also the sentence in this position could be something like 'y will be surprised', such as to invite the eventual superimposition of an outside quantifier '$\forall y$' or '$\exists y$'. To see how such a situation could arise, suppose we are given this example for analysis:

> There is someone [some cynic] who, if anyone contributes, will be surprised.

Using the letter '*y*' in our quantifier this time instead of '*x*' (just for a change), we can transcribe our example thus as a first step:

$$\exists y(\text{if anyone contributes, } y \text{ will be surprised}).$$

But the part 'if anyone contributes, *y* will be surprised' has the form '$\forall x(Fx \rightarrow p)$', or '$\forall x(Fx \rightarrow Gy)$. The whole thus becomes:

(11) $$\exists y \forall x(x \text{ contributes} \rightarrow y \text{ will be surprised}),$$

an instance of the schema '$\exists y \forall x(Fx \rightarrow Gy)$'.

The alphabetical choice of the letter of quantification is obviously arbitrary; '$\exists x\, Hx$' and '$\exists y\, Hy$' both say simply that something is an *H*. The above example of nested quantification shows, however, why we want a variety of letters to choose from; we need to keep our cross-references straight. We want it to be clear in '$\exists y \forall x(Fx \rightarrow Gy)$' that the first quantifier relates to the consequent '*Gy*' while the second relates to the antecedent '*Fx*'.

Expressions such as '*x* contributes', '*y* will be surprised', and '*x* contributes $\rightarrow$ *y* will be surprised', which are like statements except for containing '*x*' or '*y*' without a quantifier, are *open sentences;* see Chapter 21. Open sentences are neither true nor false, but they may, like terms, be said to be *true of* and *false of* various objects. The open sentence '*x* is a book' may, like the term 'book' itself, be said to be true of each book and false of everything else; and '*x* is a book . *x* is boring' may be said to be true of each boring book and false of everything else. '$x = x$' and '*x* is a man $\rightarrow$ *x* is mortal' are true of everything. In general, to say that an open sentence is true of a given object is to say that the open sentence becomes a true statement when '*x*' is reinterpreted as a name of that object. The notion of the *extension* of a term (Chapter 14) likewise carries over to open sentences: the extension of an open sentence is the class of all the objects of which the open sentence is true.

Since the schematic letters '*p*', '*q*', etc. have come now to stand as well for open sentences, e.g., '*y* will be surprised', as for statements, I shall take to calling them *sentence letters*. Note that these and the schematic term letters '*F*', '*G*', etc. differ basically in function from '*x*', '*y*', etc. Whereas '*x*' can appear in sentences—even in closed sentences, with help of a quantifier—on the other hand '*p*', '*q*', etc. and '*F*', '*G*',

etc. cannot appear in sentences at all;[3] they are merely dummy sentences and dummy terms, used in schemata which depict outward forms of sentences.

The letters 'x', 'y', 'z', and others so used, are variables. Care must be taken, however, to divorce this traditional word of mathematics from its archaic connotations. The variable is not best thought of as somehow varying through time, and causing the sentence in which it occurs to vary with it. Neither is it to be thought of as an unknown quantity, discoverable by solving equations. The variables remain mere pronouns for cross-reference; just as 'x' in its recurrences can usually be rendered 'it' in verbal translations, so the distinctive variables 'x', 'y', 'z', etc., correspond to the distinctive pronouns 'former' and 'latter', or 'first', 'second', and 'third', etc. The statement:

$$\exists x[\text{Sadie stole } x \text{ at the Emporium} \,.\, \exists y(\text{Sadie exchanged } x \text{ for } y)]$$

corresponds fairly literally to the words:

> There is something such that Sadie stole it at the Emporium and such that there is something such that Sadie exchanged the former for the latter.

In the sentence:

$$(12) \qquad \exists x(y \text{ is uncle of } x)$$

the occurrence of 'y' is free, there being no '$\forall y$' or '$\exists y$' present; but the occurrences of 'x' are bound, because of '$\exists x$'. One and the same occurrence of 'x' may be bound in a whole sentence and free in a part; the final occurrence of 'x' in (12), e.g., is bound in (12) but free in the part 'y is uncle of x'. In one and the same sentence, moreover, one occurrence of 'x' may be free and others bound; this happens in the conjunction:

$$(13) \qquad x \text{ is red} \,.\, \forall x(x \text{ has mass}),$$

in which the quantifier has to do only with the second clause. (13) means 'x is red and everything has mass', and could just as well be written with distinct variables:

[3] Exception: they may appear within quotation marks in sentences. But even a meaningless mark may appear within quotation marks in a sentence. The quotation as a whole is a meaningful name of the meaningless mark.

$$x \text{ is red } . \forall y(y \text{ has mass}).$$

By starting with 'p', 'q', 'Fx', 'Gx', 'Fy', etc., and applying quantifiers and the truth-functional notations, we obtain the *monadic quantificational schemata*. The restriction 'monadic' is due to the fact that we have not yet admitted elements like 'Fxy'. Quantificational schemata, like sentences, will be called *open* if they contain one or more free variables, and otherwise *closed*. Thus the schema:

$$\forall x[Fx \rightarrow \exists y(Gy \,.\, Hx)]$$

is closed, but its parts:

$$Fx \rightarrow \exists y(Gy \,.\, Hx), \qquad \exists y(Gy \,.\, Hx), \qquad Gy \,.\, Hx$$

are all open. Truth-functional schemata, e.g., '$p \rightarrow q$', count as closed quantificational schemata.

A quantificational schema, open or closed, is valid if it comes out true under all interpretations in all nonempty universes. A term letter is interpreted, as before, by settling what things in the universe it is to be true of. A sentence letter is interpreted, as before, by assigning it a truth value. A free variable, finally, is interpreted by assigning it some object in the universe. The concepts of consistency, implication, and equivalence extend in similar fashion.

HISTORICAL NOTE: The rules of passage were explicit in Whitehead and Russell, 1910. The term is from Herbrand. The term 'open sentence' has been used by Carnap and others. The older term for the purpose is 'propositional function', but this can be misleading, since a function in the mathematical sense is best seen as a certain type of relation rather than as a notation.

EXERCISES

1. Supposing the universe limited to $a, b, \ldots, h$, expand (7) and (8) by expanding the quantifications into alternation and conjunction.

2. There are no rules of passage for '↔'. Show, on the contrary, that no two of the schemata:

$$\forall x(p \leftrightarrow Fx), \qquad p \leftrightarrow \forall x\, Fx, \qquad p \leftrightarrow \exists x\, Fx, \qquad \exists x(p \leftrightarrow Fx)$$

are equivalent. Method: See what comes of resolution when '⊤' is put for '*p*' throughout, and again when '⊥' is put for it.

3. What implications hold between the four schemata in Exercise 2? Method: Same.

24
PRENEXITY AND PURITY

The rules of passage work in either of two directions: to widen the scope of a quantifier or to narrow it. The one direction ends by bringing the formula into *prenex* form, where all quantifiers stand in an initial row governing all the rest of the formula. The other direction ends by *purifying* the formula, by causing the scope of each quantifier to be a truth function only of components each of which shows free occurrences of the variable of the quantifier.

Both of these uses of the rules of passage depend on preparatory transformations of other kinds. If the chosen direction is the one that leads to prenex form, then certainly we must paraphrase away any occurrence of '↔' that joins formulas having quantifiers in either of them. For we saw in an exercise that there are no rules of passage for '↔'. Any quantifier that is trapped in either of the formulas joined by '↔' will remain trapped and incapable of being brought into prenex position until we get rid of '↔' by paraphrase.

Another transformation preparatory to applying the rules of passage in a prenexing way is the relettering of bound variables. '$Fx \,.\, \exists x\, Gx$' is all right as it stands, simply as a conjunction of the open schema 'Fx' and the closed schema 'there are G' (which could as well be written '$\exists y\, Gy$').

The whole says that x is an F and there are G. But it will not do to apply the rule of passage (1) of the preceding chapter to '$Fx \,.\, \exists x\, Gx$' to obtain '$\exists x(Fx \,.\, Gx)$'. This last could mean that something is round and square, which is false, while the former means that x is round and something is square, which is true of any baseball x. The thing to do before applying the rule of passage (1) to '$Fx \,.\, \exists x\, Gx$' is to rewrite this as '$Fx \,.\, \exists y\, Gy$'; then we can proceed to '$\exists y(Fx \,.\, Gy)$'. We must remember that (1) simply does not directly apply to '$Fx \,.\, \exists x\, Gx$'; the 'p' in (1) stands necessarily for a formula without free 'x'.

Once these preparations are made, we convert a schema into prenex form simply by a succession of transformations according to the rules of passage (1)–(10). Example:

$$p \leftrightarrow \forall x[Fx \rightarrow \exists y(Fy \,.\, Gx)].$$

Paraphrasing the '$\leftrightarrow$' into a conjunction of conditionals, we get:

$$p \rightarrow \forall x[Fx \rightarrow \exists y(Fy \,.\, Gx)] \,.\, \forall x[Fx \rightarrow \exists y(Fy \,.\, Gx)] \rightarrow p.$$

Relettering,

$$p \rightarrow \forall x[Fx \rightarrow \exists y(Fy \,.\, Gx)] \,.\, \forall z[Fz \rightarrow \exists w(Fw \,.\, Gz)] \rightarrow p.$$

The remaining transformations are by the rules of passage (6), (7), (2), (1), (5), and (8), two steps at a time.

$$\forall x[p \rightarrow .\, Fx \rightarrow \exists y(Fy \,.\, Gx)] \,.\, \exists z[Fz \rightarrow \exists w(Fw \,.\, Gz) .\rightarrow p].$$
$$\forall x\, \exists z[p \rightarrow .\, Fx \rightarrow \exists y(Fy \,.\, Gx) : Fz \rightarrow \exists w(Fw \,.\, Gz) .\rightarrow p].$$
$$\forall x\, \exists z[p \rightarrow \exists y(Fx \rightarrow .\, Fy \,.\, Gx) \,.\, \exists w(Fz \rightarrow .\, Fw \,.\, Gz) \rightarrow p].$$
$$\forall x\, \exists z[\exists y(p \rightarrow : Fx \rightarrow .\, Fy \,.\, Gx) \,.\, \forall w(Fz \rightarrow .\, Fw \,.\, Gz :\rightarrow p)].$$
$$\forall x\, \exists z\, \exists y \forall w(p \rightarrow : Fx \rightarrow .\, Fy \,.\, Gx :.\, Fz \rightarrow .\, Fw \,.\, Gz :\rightarrow p).$$

Note that at almost every stage there is some freedom of choice as to which quantifier to bring forward next.

The distributivity laws (9) and (10) of Chapter 22 can conduce to speed in prenexing and to brevity in the result. Consider the example:

(1) $$\forall x[\exists y(Fx \leftrightarrow Gy) \vee \exists y\, Fy] \,.\, \forall x(Fx \vee Gx).$$

According to the plan just now laid, we would begin to prenex this example by relettering thus:

(2) $$\forall x[\exists y(Fx \leftrightarrow Gy) \vee \exists z\, Fz] \,.\, \forall w(Fw \vee Gw).$$

From this result by six applications of the rules of passage (1)–(3) of Chapter 23 we would move finally to the prenex schema:

(3) $$\forall x\, \exists y\, \exists z \forall w(Fx \leftrightarrow Gy \mathbin{.\vee} Fz : Fw \vee Gw).$$

But if instead of relettering (1) as (2) we apply the distributivity laws (9) and (10) of Chapter 22 to (1), we get this:

$$\forall x[\exists y(Fx \leftrightarrow Gy \mathbin{.\vee} Fy) \,.\, Fx \vee Gx].$$

A single application of the rule of passage (1) of Chapter 23 turns this into the prenex schema:

(4) $$\forall x\, \exists y(Fx \leftrightarrow Gy \mathbin{.\vee} Fy : Fx \vee Gx),$$

which is simpler than the previous prenex result (3) and more quickly reached as well. For the sake of such benefits we can be well advised even to reletter quantifications to create duplications rather than to eliminate them. If we had been given (2) to begin with, we would have gained by relettering it as (1) and then proceeding as just now observed. Note, however, that this strategy of distributivity is the counsel only of efficiency; schemata *can* always be prenexed without it. What we strictly need to do is just paraphrase any biconditionals that contain quantifications, diversify all variables of quantification, and apply the rules of passage.

In relettering it is important always to avoid collisions of variables. It is all right to change '$\exists y(Fx \vee Gy)$' to '$\exists z(Fx \vee Gz)$', but decidedly wrong to change it to '$\exists x(Fx \vee Gx)$'; for this, far from being a superficial alphabetical change of notation, is a crucial change of structure. When a quantification is relettered, the schema that stands as the scope of the quantifier must come to exhibit free occurrences of the new variable at all *and only* the places where it had exhibited free occurrences of the old variable. The '$Fx \vee Gz$' of the example above has free 'z' where and only where '$Fx \vee Gy$' had free 'y'; the '$Fx \vee Gx$', on the other hand, has free 'x' where but *not* only where '$Fx \vee Gy$' had free 'y'.

Whereas prenexing brings quantifiers out, purification drives them in. Purification, like prenexing, rests on preparatory transformations. Thus consider the impure quantification '$\exists x(Fx \,.\, p \vee Gx)$', whose impurity is '$p$', lacking '$x$'. As a preparatory transformation we convert the scope '$Fx \,.\, p \vee Gx$' of the quantifier into alternational normal form and then distribute the quantifier through the alternation, thus:

$$\exists x(Fx \,.\, p \,.\vee.\, Fx \,.\, Gx),$$
$$\exists x(Fx \,.\, p) \vee \exists x(Fx \,.\, Gx).$$

Now we can apply the rule of passage (1) of Chapter 23, completing the purification.

$$\exists x\, Fx \,.\, p \,.\vee\, \exists x(Fx \,.\, Gx).$$

To purify a universal quantification whose impurity was similarly buried, we would begin by converting the scope of the quantifier to conjunctional normal form; for a universal quantifier distributes through conjunction. E.g., by steps dual to those above, the impure quantification '$\forall x(Fx \vee.\, p \,.\, Gx)$' fares as follows:

$$\forall x(Fx \vee p \,.\, Fx \vee Gx),$$
$$\forall x(Fx \vee p) \,.\, \forall x(Fx \vee Gx),$$
$$\forall x\, Fx \vee p \,.\, \forall x(Fx \vee Gx).$$

By a combination of these methods we can rid any quantificational schema of all impurities. The general routine is as follows. Wherever there is an impure quantification, put the scope of the quantifier into alternational or conjunctional normal form (according as the quantifier was existential or universal). However, prior to this and at every stage of the process, apply any applicable one of the rules of passage (1)–(8) of Chapter 23 on sight—always so as to drive quantifiers inward. Continue until all impurities are gone.

E.g., let us purify (1). By the rule of passage (4) of Chapter 23, (1) becomes:

(5) $$\forall x\, \exists y(Fx \leftrightarrow Gy) \vee \exists y\, Fy \,.\, \forall x(Fx \vee Gx).$$

Next we turn the '$Fx \leftrightarrow Gy$' of the impure quantification '$\exists y(Fx \leftrightarrow Gy)$' into alternational normal form and distribute the '$\exists y$' through it. The whole of (5) becomes:

$$\forall x[\exists y(Fx \,.\, Gy) \vee \exists y(-Fx \,.\, -Gy)] \vee \exists y\, Fy \,.\, \forall x(Fx \vee Gx).$$

Now we apply the rule of passage (1) of Chapter 23 twice, getting:

$$\forall x[Fx \,.\, \exists y\, Gy \,.\vee.\, -Fx \,.\, \exists y\, -Gy] \vee \exists y\, Fy \,.\, \forall x(Fx \vee Gx).$$

Next we turn the scope of the initial '$\forall x$' into conjunctional normal form. The whole becomes:

$$\forall x[Fx \vee \exists y\, -Gy \,.\, \exists y\, Gy \vee -Fx \,.\, \exists y\, Gy \vee \exists y\, -Gy] \\ \vee \exists y\, Fy \,.\, \forall x(Fx \vee Gx).$$

Now we can apply the rule of passage (2) of Chapter 23.

$$\forall x[Fx \vee \exists y\, -Gy \,.\, \exists y\, Gy \vee -Fx] \,.\, \exists y\, Gy \vee \exists y\, -Gy \\ .\vee \exists y\, Fy : \forall x(Fx \vee Gx).$$

Next we distribute the '$\forall x$'.

$$\forall x(Fx \vee \exists y\, -Gy) \,.\, \forall x(\exists y\, Gy \vee -Fx) \,.\, \exists y\, Gy \vee \exists y\, -Gy \\ .\vee \exists y\, Fy : \forall x(Fx \vee Gx).$$

Now we can apply the rule of passage (4) of Chapter 23 twice.

(6) $$\forall x\, Fx \vee \exists y\, -Gy \,.\, \exists y\, Gy \vee \forall x\, -Fx \,.\, \exists y\, Gy \vee \exists y\, -Gy \\ .\vee \exists y\, Fy : \forall x(Fx \vee Gx).$$

This at last is pure. If we transcribe '$\forall x$' as '$-\exists x-$' and then drop all variables, we can condense the whole into the Boolean schema:

(7) $$-\exists \bar{F} \vee \exists \bar{G} \,.\, \exists G \vee -\exists F \,.\, \exists G \vee \exists \bar{G} \,.\vee\, \exists F : -\exists -(F \vee G).$$

Thus the original schema (1), the prenex schemata (3) and (4), the pure schema (6), and the Boolean schema (7) all are equivalent.

An alternative method of purification is also worth noting. Here the quantifier may be universal or existential indifferently; let us represent it as 'Qx'. The impurity, let us say, is 'p'; so the quantification may be pictured as '$Qx(\ldots p\ldots)$'. Now this is equivalent to:

(8) $$p \,.\, Qx(\ldots \top \ldots) \,.\vee.\, \bar{p} \,.\, Qx(\ldots \bot \ldots)$$

—let the reader reason why. Finally we resolve '⊤' and '⊥' out of '. . . ⊤ . . .' and '. . . ⊥ . . .' by truth-function logic. The instructions, we see, are admirably brief. The reader would do well to run some complex examples through by both methods and compare them for efficiency.

A monadic quantificational schema has impurities as long as it has stacked quantifiers, i.e., as long as any quantifier has a quantifier in its scope. For, if a schema has any stacked quantifiers, then it will have at least two innermost stacked quantifiers—say '$\exists x$' and '$\forall y$'. By this I mean that the scope of '$\exists x$' is a truth function of schemata one of which is the '$\forall y$' quantification, and the '$\forall y$' quantification contains no further quantifiers. But then either the '$\exists x$' quantification or the '$\forall y$' quantification is bound to be impure. For, if the '$\forall y$' quantification lacks free 'x', it renders the '$\exists x$' quantification impure. If on the other hand it has free 'x', then the scope of '$\forall y$' is a truth function of schemata one of which is 'Fx' or 'Gx' or the like; and this renders the '$\forall y$' quantification impure.

We have seen how to purify any monadic quantificational schema, and now we have seen further that the result will be devoid of stacked quantifiers. Each quantification will have as its scope a truth function merely of 'Fx', 'Gx', etc., if 'x' is the variable of the quantifier; 'Fy', 'p', etc. are driven out by the purification, and likewise all stacked quantifiers. Then, if we transcribe '$\forall x$' again as '$- \exists x -$' and then drop the variables, as we did in passing from (6) to (7), all quantifications will disappear in favor of Boolean existence schemata.

We cannot quite say that every monadic quantificational schema is thus reducible to a Boolean schema, such as (7). (6) was thus reducible because it contained no sentence letters 'p', 'q', etc., and no unquantified 'Fz' or the like. But we can say that every monadic schema that is *closed* (and hence devoid of unquantified 'Fz' and the like) is reducible in the observed fashion to a truth function of Boolean existence schemata and sentence letters. The method consists in purifying and finally transcribing universal quantifiers and dropping variables.

It thus emerges that the notational apparatus of monadic quantificational schemata, with its variables of quantification and all its stacked quantifiers and impurities, is needlessly inflated. The letters and truth functions of Part I and the Boolean existence schemata would have been apparatus enough. Still, there is an important reason for taking up the inflated apparatus: it prepares us for the advent in Part III of polyadic schemata, containing 'Fxy' and the like. At that point the kind of reducibility here observed will lapse.

HISTORICAL NOTE: Two names are particularly associated respectively with prenexing and purifying: those of Skolem and Behmann. Skolem depended heavily on the prenex form for his proof procedure in polyadic quantification theory, much as we shall do. On the other hand Behmann's decision procedure of 1922 (see end of Chapter 19 above) depended on driving quantifiers in.

EXERCISES

1. Put the schema '$\forall x\, Fx \rightarrow .\ \forall x\, Gx \leftrightarrow \forall x\, Hx$' into prenex form.
2. Likewise '$\exists x\, Fx \leftrightarrow \exists x\, Gx \mathrel{.\!\rightarrow} \exists x\, Hx$'.
3. Purify '$\exists x \forall y(Fx \leftrightarrow Gy \mathrel{.\!\rightarrow} Hy)$' by the first method.
4. Likewise '$\forall x \exists y(Fx \rightarrow Gy \mathrel{.\!\leftrightarrow} Hy)$'.
5. Repeat 3 and 4 by the second method.
6. Establish by informal argument the equivalence of (8) to '$Qx(\ldots p \ldots)$'.

25
VALIDITY AGAIN

A universal quantifier is indifferent to the validity of the schema that it governs. This is evident from the meaning of universal quantification. To say e.g. that '$\forall y(Fy \rightarrow \exists x\, Fx)$' comes out true under all interpretations of 'F' is the same as saying that '$Fy \rightarrow \exists x\, Fx$' comes out true under all interpretations of 'F' and all assignments of objects to 'y'; and such is the meaning of validity of '$Fy \rightarrow \exists x\, Fx$'. (See Chapter 23.) Correspondingly an existential quantifier is indifferent to the consistency of a schema that it governs.

We can test an open schema for validity, then, by testing its *universal closure*. This is a closed schema that we get from it by prefixing a universal quantifier for each free variable. In order to test a closed schema

for validity, so long as it has no sentence letters, we reduce it to a Boolean schema and then apply the method of existential conditionals or the cellular method (Chapter 19) as we please. How to reduce it to a Boolean schema was illustrated in Chapter 24 in the transformation of (1) progressively into (5), (6), and finally (7).

Let us test the open schemata:

$$\forall x\, Fx \to Fy, \qquad Fy \to \exists x\, Fx$$

for validity. As remarked, this means testing the closed schemata:

$$\forall y(\forall x\, Fx \to Fy), \qquad \forall y(Fy \to \exists x\, Fx).$$

Purifying them by means of the rules of passage (6) and (8) of Chapter 23, we get:

$$\forall x\, Fx \to \forall y\, Fy, \qquad \exists y\, Fy \to \exists x\, Fx.$$

If despite their patent validity we persist in reducing these to Boolean schemata, we get:

$$-\exists \bar{F} \to -\exists \bar{F}, \qquad \exists F \to \exists F.$$

For a less trivial example, consider the inference:

$$Fz, \qquad \forall x(\exists y\, Fy \to Gx), \qquad \therefore \exists y(Fy \,.\, Gy).$$

To check the implication we test the open conditional:

$$Fz \,.\, \forall x(\exists y\, Fy \to Gx) \mathrel{.\!\!\to} \exists y(Fy \,.\, Gy)$$

for validity. This means testing the closed schema:

$$\forall z[Fz \,.\, \forall x(\exists y\, Fy \to Gx) \mathrel{.\!\!\to} \exists y(Fy \,.\, Gy)].$$

Proceeding to purify it, we apply successively the rules of passage (8), (1), and (6) of Chapter 23.

$$\exists z[Fz \,.\, \forall x(\exists y\, Fy \to Gx)] \to \exists y(Fy \,.\, Gy),$$
$$\exists z\, Fz \,.\, \forall x(\exists y\, Fy \to Gx) \mathrel{.\!\!\to} \exists y(Fy \,.\, Gy),$$
$$\exists z\, Fz \,.\, \exists y\, Fy \to \forall x\, Gx \mathrel{.\!\!\to} \exists y(Fy \,.\, Gy).$$

Then we transcribe the universal quantifier and drop the variables.

$$\exists F \,.\, \exists F \rightarrow - \exists \bar{G} \,.\rightarrow \exists FG.$$

Choosing now the method of existential conditionals, we turn this into conjunctional normal form and simplify. It boils down to:

$$- \exists F \vee \exists \bar{G} \vee \exists FG.$$

Collecting the affirmative existential schemata as '$\exists(\bar{G} \vee FG)$', we end up with the existential conditional:

$$\exists F \rightarrow \exists(\bar{G} \vee FG).$$

We check its validity by criterion (iv) of Chapter 19. That is, we check by fell swoop that 'F' implies '$\bar{G} \vee FG$'.

Note that this technique, again, depends for its soundness on a law of interchange of equivalents. Again, however, as in the case of truth-functional schemata (Chapter 9) and Boolean schemata (Chapter 19), all is in order; for it is evident here again that the truth value of a compound statement depends on no features of the component sentences and terms except their being true or false or true of something or of nothing.

We turn finally to the question of a validity test for closed monadic quantificational schemata containing sentence letters. Such a schema will be valid, clearly, just in case it resolves to '$\top$' or to a valid schema under each substitution of '$\top$' and '$\bot$' for its sentence letters. So the test is truth-value analysis. If the analysis issues in '$\bot$' at any point, of course the schema was not valid. Otherwise we follow up with validity tests of any quantificational schemata issuing from the truth-value analysis.

Example:

$$\forall x(Fx \leftrightarrow p) \,.\, \exists x\, Fx \leftrightarrow q \,.\rightarrow.\, p \leftrightarrow q$$

left side:

$$\forall x(Fx \leftrightarrow \top) \,.\, \exists x\, Fx \leftrightarrow q \,.\rightarrow.\, \top \leftrightarrow q$$

$$\forall x\, Fx \,.\, \exists x\, Fx \leftrightarrow q \,.\rightarrow q$$

$$\forall x\, Fx \,.\, \exists x\, Fx \leftrightarrow \top \,.\rightarrow \top \qquad \forall x\, Fx \,.\, \exists x\, Fx \leftrightarrow \bot \,.\rightarrow \bot$$

$$\top \qquad\qquad -(\forall x\, Fx \,.\, - \exists x\, Fx)$$

right side:

$$\forall x(Fx \leftrightarrow \bot) \,.\, \exists x\, Fx \leftrightarrow q \,.\rightarrow.\, \bot \leftrightarrow q$$

$$\forall x\, -Fx \,.\, \exists x\, Fx \leftrightarrow q \,.\rightarrow \bar{q}$$

$$\forall x\, -Fx \,.\, \exists x\, Fx \leftrightarrow \top \,.\rightarrow \bot \qquad \forall x\, -Fx \,.\, \exists x\, Fx \leftrightarrow \bot \,.\rightarrow \top$$

$$-(\forall x\, -Fx \,.\, \exists x\, Fx) \qquad\qquad \top$$

The question of validity here reduces to the question of validity of '$-(\forall x\ Fx\ .\ -\exists x\ Fx)$' and '$-(\forall x\ -Fx\ .\ \exists x\ Fx)$'. These go over into Boolean schemata as '$-(-\exists \bar{F}\ .\ -\exists F)$' and '$-(-\exists F\ .\ \exists F)$'. The latter has the valid form '$-(\bar{p}p)$'. The other goes into conjunctional normal form as '$\exists \bar{F} \vee \exists F$', and so, by collection of affirmative existence schemata, '$\exists(\bar{F} \vee F)$'. This is valid by criterion (i) of Chapter 19.

Dismissal of '$-(-\exists F\ .\ \exists F)$' as of the form '$-(\bar{p}p)$' was an obvious shortcut, falling outside the mechanical routine thus far set down. That routine would prescribe converting '$-(-\exists F\ .\ \exists F)$' to conjunctional normal form as '$\exists F \vee -\exists F$', then rendering this as an existential conditional '$\exists F \rightarrow \exists F$', and finally observing that the term schema in the antecedent implies the one in the consequent by fell swoop. It is worth noting that the mechanical routine covers such cases, but it would be silly to follow it.

When truth-value analysis is applied to quantificational schemata, as above, four supplementary rules of resolution are sometimes needed in addition to those in Chapter 5. Namely, '$\forall x\top$', '$\exists x\top$', '$\forall x\bot$', and '$\exists x\bot$' resolve to '$\top$', '$\top$', '$\bot$', and '$\bot$'. One of these is used when we test the following rule of passage, (8) of Chapter 23.

$$\forall x(Fx \rightarrow p) \leftrightarrow .\ \exists x\ Fx \rightarrow p$$

$$\forall x(Fx \rightarrow \top) \leftrightarrow .\ \exists x\ Fx \rightarrow \top \qquad \forall x(Fx \rightarrow \bot) \leftrightarrow .\ \exists x\ Fx \rightarrow \bot$$

$$\forall x\top \qquad\qquad \forall x\ -Fx \leftrightarrow -\exists x\ Fx$$

$$\top$$

The remainder of this test consists in testing the leftover biconditional '$\forall x\ -Fx \leftrightarrow -\exists x\ Fx$' for validity. It goes over into the truth-functionally valid Boolean schema '$-\exists F \leftrightarrow -\exists F$'.

We conclude with a full-dress example, starting in words.

PREMISES: If the Bissagos report is to be trusted then the *chargé d'affaires* is a mere tool of the sisal interests and none of the natives really favored the coupon plan.

If the *chargé d'affaires* is a mere tool of the sisal interests then some of the natives either really favored the coupon plan or were actuated by a personal animosity against the deputy resident.

CONCLUSION: If the Bissagos report is to be trusted then some who were actuated by a personal animosity against the deputy resident did not really favor the coupon plan.

Putting 'p' for 'the Bissagos report is to be trusted', 'q' for 'the *chargé d'affaires* is a mere tool of the sisal interests', 'Fx' for 'x is a native', 'Gx' for 'x really favored the coupon plan', and 'Hx' for 'x was actuated by a personal animosity against the deputy resident', we can represent the premises and conclusion by the mixed schemata:

$$p \rightarrow .\ q\ .\ \forall x(Fx \rightarrow -Gx),$$
$$q \rightarrow \exists x(Fx\ .\ Gx \vee Hx),$$
$$p \rightarrow \exists x(Hx\ .\ -Gx).$$

Accordingly we submit the conditional:

$$p \rightarrow .\ q\ .\ \forall x(Fx \rightarrow -Gx) : q \rightarrow \exists x(Fx\ .\ Gx \vee Hx) \mathrel{:\!\!\rightarrow}. p \rightarrow \exists x(Hx\ .\ -Gx)$$

to truth-value analysis. Let us save space by abbreviating the three quantifications as 'A', 'B', and 'C'.

$$\frac{\dfrac{p \rightarrow qA\ .\ q \rightarrow B \mathrel{.\!\!\rightarrow}. p \rightarrow C}{\dfrac{\mathsf{T} \rightarrow qA\ .\ q \rightarrow B \mathrel{.\!\!\rightarrow}. \mathsf{T} \rightarrow C}{qA\ .\ q \rightarrow B \mathrel{.\!\!\rightarrow} C}}{\dfrac{\mathsf{T}A\ .\ \mathsf{T} \rightarrow B \mathrel{.\!\!\rightarrow} C}{AB \rightarrow C}} \qquad \frac{\dfrac{\perp \rightarrow qA\ .\ q \rightarrow B \mathrel{.\!\!\rightarrow}. \perp \rightarrow C}{\mathsf{T}}}{\dfrac{\perp A\ .\ \perp \rightarrow B \mathrel{.\!\!\rightarrow} C}{\mathsf{T}}}$$

So our problem reduces to a validity test of '$AB \rightarrow C$', which is:

$$\forall x(Fx \rightarrow -Gx)\ .\ \exists x(Fx\ .\ Gx \vee Hx) \mathrel{.\!\!\rightarrow} \exists x(Hx\ .\ -Gx).$$

Transcribed as a Boolean schema, this becomes:

$$-\exists -(F \rightarrow \bar{G})\ .\ \exists[F(G \vee H)] \mathrel{.\!\!\rightarrow} \exists H\bar{G},$$

or, in conjunctional normal form,

$$\exists -(F \rightarrow \bar{G}) \vee -\exists[F(G \vee H)] \vee \exists H\bar{G}.$$

When we collect the affirmative existence schemata, this becomes:

$$-\exists[F(G \vee H)] \vee \exists[-(F \rightarrow \bar{G}) \vee H\bar{G}]$$

and so the existential conditional:

$$\exists[F(G \vee H)] \rightarrow \exists[-(F \rightarrow \bar{G}) \vee H\bar{G}].$$

Finally we check the validity of this, according to the criterion (iv) of Chapter 19, by verifying that '$F(G \vee H)$' implies '$-(F \rightarrow \bar{G}) \vee H\bar{G}$'. This again is done by truth-value analysis.

$$F(G \vee H) \rightarrow [-(F \rightarrow \bar{G}) \vee H\bar{G}]$$

$$F(\top \vee H) \rightarrow [-(F \rightarrow \bot) \vee H\bot] \qquad F(\bot \vee H) \rightarrow [-(F \rightarrow \top) \vee H\top]$$

$$F \rightarrow F \qquad FH \rightarrow H$$

$$\top \qquad \top$$

EXERCISES

1. Test these six schemata for validity.

$$\forall y(Fy \rightarrow \forall x\, Fx), \qquad \exists y(Fy \rightarrow \forall x\, Fx),$$
$$\forall y(\exists x\, Fx \rightarrow Fy), \qquad \exists y(\exists x\, Fx \rightarrow Fy),$$
$$\exists y(\forall x\, Fx \rightarrow Fy), \qquad \exists y(Fy \rightarrow \exists x\, Fx).$$

2. Check the biconditionals corresponding to (1)–(7) of Chapter 23.

3. Test each of these pairs for equivalence, by testing the biconditional,

$$\forall x\, Fx, \quad \forall x(Fx \vee Gx) \,.\, \forall x(Fx \vee -Gx);$$
$$\forall x\, Fx, \quad \forall x(Fx \,.\, Gx) \vee \forall x(Fx \,.\, -Gx);$$
$$\exists x\, Fx, \quad \exists x(Fx \vee Gx) \,.\, \exists x(Fx \vee -Gx);$$
$$\exists x\, Fx, \quad \exists x(Fx \,.\, Gx) \vee \exists x(Fx \,.\, -Gx).$$

4. Check the following argument.

PREMISES: The persons responsible for the recent kidnappings are experimental psychologists.
If no experimental psychologists are known to the police, then none of the former bosses of the bootleg ring are experimental psychologists.

CONCLUSION: If any of the former bosses of the bootleg ring are responsible for the recent kidnappings then some experimental psychologists are known to the police.

26
SUBSTITUTION

In logical theory the schema is what is in focus, but still it is the sentence finally that matters. The purpose of schemata is simply to facilitate the logical study of sentences by depicting their logical forms. What relates the schemata to the sentences that they schematize is substitution. We get the sentence by substituting in the schema. It is time to examine this connection.

For truth-functional schemata the relevant notion of substitution was clear and simple: just substitute sentences for sentence letters, and always the same sentence for the same letter. For Boolean schemata the notion was equally simple: just substitute general terms for term letters, and the same term for the same letter.

The move to quantification complicates substitution by allowing 'Fx' as a whole to stand for an unanalyzed sentence. Clearly we cannot speak simply of substitution of arbitrary sentences for 'Fx', 'Fy', etc.; a certain correspondence must be preserved between the substitutes for 'Fx' and 'Fy'. What correspondence?

First approximation: the substitutes must be sentences that are alike except that the one has free occurrences of 'x' in all and only the places where the other has free occurrences of 'y'. Examples:

	Fx	Fy
(1)	x is proud of the team,	y is proud of the team.
(2)	x is proud of z,	y is proud of z.
(3)	x is proud of x,	y is proud of y.
(4)	$\exists z$(x is proud of z),	$\exists z$(y is proud of z).

In each pair the one sentence says of x what the other says of y: that he is proud of the team, proud of z, proud of himself, proud of something. Still, the requirement that the free occurrences of 'x' wholly match those of 'y' is too restrictive. For, along with pride of self, pride of something, pride of the team, pride of z, another alternative that needs to be allowed is pride of x:

(5)	x is proud of x,	y is proud of x.

This is not to disallow the pair (3); it is additional. Both pairs need to be recognized as legitimate and different pairs of substitutions for 'Fx' and 'Fy'.

Thus the desired requirement, vaguely speaking, is that the sentences substituted for 'Fx' and 'Fy' be sentences that say respectively of x and y *some* one same thing—be it pride of self or pride of x.

The right device lies ready to hand, back in Chapter 21: term abstraction. We have merely to recognize 'F' separately as a term letter again. Substitution for 'F' in 'Fx' consists in two steps: in putting a suitable term abstract for 'F' in 'Fx' and then performing concretion ((6) or (13) of Chapter 21). Similarly for substitution in 'Fy'. For obtaining the pair of results (1) the appropriate abstract is '$\{w$: w is proud of the team$\}$'; for (2) it is '$\{w$: w is proud of $z\}$'; for (3) it is '$\{w$: w is proud of $w\}$'; for (4) it is '$\{w$: $\exists z(w$ is proud of $z)\}$'; and for (5) it is '$\{w$: w is proud of $x\}$'. Term abstracts figure not at all in quantificational schemata nor in the tests of validity, but they do yeoman service when we want to apply those schemata and come out with sentences. They are wanted to govern substitution.

Certain restraints are needed. Substitution of the abstract '$\{w$: $\exists x(w$ is proud of $x)\}$' for 'F' in 'Fx' and 'Fy' might seem to give the pair of sentences:

(6) $\exists x(x$ is proud of $x)$, $\exists x(y$ is proud of $x)$.

This pair, unlike (1)–(5), is ill matched. There is no one "same thing" that the left sentence says about x and the right about y. The sentence at the right says that y is proud of something, but the one at the left says nothing at all about x; it says only that someone is proud of himself, and it could as well be rendered '$\exists z(z$ is proud of $z)$'. It has no free 'x'.

Hence this *first restriction:* An abstract is not to be predicated of a variable that would be captured by a quantifier inside the abstract.

Suppose next that we are substituting for 'F' not just in 'Fx' and 'Fy' as heads of two columns (1)–(5), but in the valid schema '$\forall x\ Fx \rightarrow Fy$'. Substitution of '$\{w$: w is proud of $x\}$' for 'F' in 'Fx' and 'Fy' had been legitimate, yielding the pair of sentences (5); but it is not legitimate in the context '$\forall x\ Fx \rightarrow Fy$', for it would yield:

(7) $\forall x(x$ is proud of $x) \rightarrow .\ y$ is proud of x,

which fails to inherit the logical structure to which '$\forall x\ Fx \rightarrow Fy$' owed its validity. This last says that if everything is F then y is; (7) says that if everyone is proud of himself then, irrelevantly, this fellow is proud of that one.

Broadly speaking, the trouble in (7) is the same as in (6): an occurrence of 'x' has been captured by a quantifier. The causes, however, are opposite. In (6) a quantifier that had been lurking in the substituted abstract '$\{w$: $\exists x(w$ is proud of $x)\}$' captured the free 'x' that was waiting in 'Fx'. In (7), conversely, a quantifier lurking in the context of 'Fx' captured the free 'x' of the substituted abstract '$\{w$: w is proud of $x\}$'. So we need still this *second restriction:* A variable free in the abstract must not be such as to be captured by a quantifier in the schema into which the abstract is substituted.

We may sum up the two restrictions symmetrically thus: Quantifiers of the substituted abstract must not capture variables of the schema in which the substitution takes place, and variables of the substituted abstract must not be captured by quantifiers of the schema in which the substitution takes place. These restrictions simply ward off confusions of variables which, if allowed, would cause substitution to deviate from its intended purpose of interpreting term letters.

In turning quantificational schemata into sentences we have not only term letters but also sentence letters to reckon with. For sentences the appropriate operation of substitution is simpler, there being no question of abstracts and concretion. Substitution of a sentence for a sentence letter in a schema consists as usual in putting the sentence for all occurrences of the letter. The first of the two restrictions set forth above, moreover, no longer has a place, there being no predicating of abstracts. But the *second restriction carries over:* Variables free in the substituted sentence must not be such as to be captured by quantifiers in the schema into which the sentence is substituted. This restriction merely makes explicit the understanding that governed 'p' in Chapter 23: 'p' represented, in those contexts, a sentence devoid of free 'x'.

EXERCISES

1. Substitute each of the abstracts:

$\{z$: x praised z to $y\}$, $\{z$: z praised z to $y\}$,
$\{z$: y praised z to $z\}$, $\{z$: y praised z to $x\}$

for 'F' in 'Fx'. Compatibly with the restrictions on substitution, which of these abstracts can be substituted for 'F' in '$\exists x\, Fx$'? What do the resulting sentences mean?

2. Find an abstract which, substituted for 'F' in '$Fx \rightarrow Fy$', yields:

$$y^x = xx + x \mathrel{.\!\!\rightarrow\!\!.} y^y = yx + y.$$

III
GENERAL THEORY OF QUANTIFICATION

27
SCHEMATA EXTENDED

In the logical tradition terms are distinguished into two kinds, *relative* and *absolute*. The characteristic of a relative term is that it describes things only relatively to further things which have afterward to be specified in turn. Thus 'father' as in 'father of Isaac', and 'north' as in 'north of Boston', are relative terms. What were spoken of as terms in Chapter 14, on the other hand, are absolute terms. Words capable of behaving as relative terms can regularly be used *also* as absolute terms, through what amounts to a tacit existential quantification in the context; thus we may say absolutely that Abraham is a father, meaning that there is something of which Abraham is a father.

In English a convenient earmark of the relative use of a term is the adjoining of an 'of'-phrase or possessive modifier whose sense is not that of ownership. Thus 'father of Isaac', or 'Isaac's father', has nothing to do with proprietorship on Isaac's part, but means merely 'that which bears the father-relation to Isaac'. We can appreciate the distinction between the possessive 'my' and the relative 'my' by recalling what Dionysodorus said to Ctesippus with reference to the latter's dog: '. . . he is a father, and he is yours; therefore he is your father' (Plato, *Euthydemus*).

It will be better hereafter to speak of terms not as absolute and relative, but rather as *monadic* and *polyadic*, more particularly *dyadic*. The change in terminology reflects a change in point of view. The term 'father' as relative is thought of as referring still to Abraham, even though relatively somehow to Isaac. The term 'father' or 'father of' as dyadic is thought of rather as true of two men coordinated as an ordered pair, Abraham and Isaac.

A dyadic term, like a monadic one, may occur indifferently as substantive, adjective, or verb. In 'x is a helper of y' we use the substantive, in 'x is helpful toward y' the adjective, and in 'x helps y' the verb; but logically there is no need to distinguish the three. Logically the important thing is that whereas the monadic terms 'man', 'walks', etc. are true of Caesar, Socrates, etc. one by one, on the other hand the dyadic term 'helps' is true of Jesus and Lazarus as a pair, and of Uncle Tom and Little Eva as a pair. If as in the foregoing pages we write 'Fx' for 'x is an F',

then the analogous notation in connection with relative terms should be '*Fxy*', '*x* is *F* to *y*'.

The order of '*x*' and '*y*' in '*x* helps *y*' is in one respect accidental: '*x* helps *y*' can as well be phrased '*y* is helped by *x*'. But in another respect the order is essential: '*x* helps *y*', e.g., 'Jesus helps Lazarus', is not equivalent to '*y* helps *x*'. So the sentence '*x* helps *y*' may be described equally as of the form '*Fxy*' and as of the form '*Fyx*', but the interpretations thus successively imposed on '*F*' are then distinct from each other—as distinct as 'helps' and 'is helped by'. '*Fxy*' cannot in general be equated with '*Fyx*'.

Besides dyadic terms we may recognize *triadic* ones, *tetradic*, and so on; e.g., '*Gxyz*' may mean '*x* gives *y* to *z*', and '*Hxyzw*' may mean '*x* pays *y* to *z* for *w*'.

There are forms of inference, logically no less sound than those dealt with in Part II, which are insusceptible to the methods of Part II simply because their analysis calls for recognition of dyadic terms. An example from Jungius (fl. 1640) is:

All circles are figures; ∴ All who draw circles draw figures.

The premise can be represented in our previous notation as '$\forall x(Fx \rightarrow Gx)$', but the conclusion presents difficulties. We can indeed represent the conclusion as '$\forall x(Hx \rightarrow Jx)$', interpreting '*Hx*' as '*x* draws a circle' and '*Jx*' as '*x* draws a figure', but then the schemata '$\forall x(Fx \rightarrow Gx)$' and '$\forall x(Hx \rightarrow Jx)$' bear no visible interconnection which could justify inference of the one from the other. What we must do is extend our category of quantificational schemata to admit such forms as '*Hyx*' for '*y* draws *x*'. Then '*y* draws a circle' can be represented as '$\exists x(Fx \,.\, Hyx)$', and '*y* draws a figure' as '$\exists x(Gx \,.\, Hyx)$'; thereupon our conclusion as a whole, 'All who draw circles draw figures', becomes:

(1) $$\forall y[\exists x(Fx \,.\, Hyx) \rightarrow \exists x(Gx \,.\, Hyx)].$$

Quantification theory needs to be extended in such a way as to enable us to show, among other things, that '$\forall x(Fx \rightarrow Gx)$' implies (1).

Another example of the need of thus extending quantification theory is this:

PREMISE: There is a painting that all critics admire;
CONCLUSION: Every critic admires some painting or other.

With 'Gx' interpreted as 'x is a critic', and 'Hxy' as 'x admires y', we may represent 'all critics admire y' as '$\forall x(Gx \rightarrow Hxy)$'. So, interpreting '$Fy$' as '$y$ is a painting', we may represent the above premise as:

(2) $$\exists y[Fy \,.\, \forall x(Gx \rightarrow Hxy)].$$

Further, since 'x admires some painting or other' becomes '$\exists y(Fy \,.\, Hxy)$', the conclusion as a whole takes on the form:

(3) $$\forall x[Gx \rightarrow \exists y(Fy \,.\, Hxy)].$$

One more example:

PREMISE: There is a philosopher whom all philosophers contradict,

CONCLUSION: There is a philosopher who contradicts himself.

The premise here has a form closely similar to (2).

(4) $$\exists y[Fy \,.\, \forall x(Fx \rightarrow Gxy)].$$

The conclusion is simply '$\exists x(Fx \,.\, Gxx)$'.

We saw in Chapter 22 how differences in grouping could affect the meaning of a quantification; '$\forall x(Fx \vee Gx)$' had to be distinguished from '$\forall x\, Fx \vee \forall x\, Gx$', and '$\exists x(Fx \,.\, Gx)$' from '$\exists x\, Fx \,.\, \exists x\, Gx$'. Considerations of this kind come to loom larger now that we need quantifications within quantifications. Thus, let us reflect next on the expression:

(5) $$\forall x[Fx \rightarrow \exists y(Fy \,.\, Gxy)].$$

If we interpret 'Fx' as 'x is a number' and 'Gxy' as 'x is less than y', then (5) comes to mean:

Every number is such that some number exceeds it,

or briefly 'Every number is exceeded by some number'. This might carelessly be rephrased 'Some number exceeds every number' and then be put back into symbols as:

$\exists y(y$ is a number $.\, y$ exceeds every number),

i.e., (4). But actually there is all the difference between (5) and (4) that there is between truth and falsity. (5) says that for every number there is a larger, which is true, whereas (4) says there is some great number which, at once, exceeds every number. This last is false on two counts: for there is no greatest number, and even if there were it would not exceed itself.

The distinction in form between (3) and (2) is the same as just now stressed between (5) and (4). The wording of the premise and conclusion about paintings illustrates again the awkwardness of ordinary language in keeping the distinction clear. The notation of quantification is handier in this respect.

The mathematical concept of limit provides, for readers familiar with it, an apt further illustration of the above distinction. A function $f(x)$ is said to approach a limit h, as x approaches k, if for every positive number ε there is a positive number δ such that $f(x)$ is within ε of h for every x ($\neq k$) within δ of k. In terms of quantifiers this condition appears as follows:

$$\forall \varepsilon\{\varepsilon > 0 .\!\rightarrow \exists \delta[\delta > 0 \,.\, \forall x(0 < |x - k| < \delta .\!\rightarrow\!. |f(x) - h| < \varepsilon)]\}.$$

As textbooks rightly emphasize, we must think of ε as chosen first; for each choice of ε a suitable δ can be chosen. This warning is, in effect, a warning against confusing the above formula with the essentially different one:

$$\exists \delta\{\delta > 0 \,.\, \forall \varepsilon[\varepsilon > 0 .\!\rightarrow \forall x(0 < |x - k| < \delta .\!\rightarrow\!. |f(x) - h| < \varepsilon)]\}.$$

The distinction between these two formulas will be recognized as identical with that between (5) and (4), and between (3) and (2).

The essential contrast between (5) and (4), and between (3) and (2), becomes simpler and more striking when we compare:

(6) $$\forall x\ \exists y\ Fxy, \qquad \exists y\ \forall x\ Fxy.$$

Suppose we interpret 'Fxy' as 'x and y are the same thing', so that the schemata (6) become:

(7) $\forall x \exists y$(x and y are the same thing),
(8) $\exists y \forall x$(x and y are the same thing).

For each chosen object x, clearly there will be an object which is the same (viz., the chosen object x itself). Of each object x, therefore, the sentence:

$$\exists y(x \text{ and } y \text{ are the same thing})$$

is true. So (7) is true. On the other hand, as long as there are more objects than one in the universe, no one object can be the same as each; i.e., no one object y can be such that

$$\forall x(x \text{ and } y \text{ are the same thing}).$$

So (8) is false.

In general '$\forall x\ \exists y\ Fxy$' says that once any object whatever x is fixed upon, an object y is forthcoming such that Fxy. Different choices of x may bring forth different choices of y. On the other hand '$\exists y\ \forall x\ Fxy$' says that some object y can be fixed upon such that, for this same fixed y, 'Fxy' will hold for all comers x.

Supposing a limited universe of objects $a, b, \ldots, h$, let us see how '$\forall x\ \exists y\ Fxy$' and '$\exists y\ \forall x\ Fxy$' compare when the quantifications are expanded into conjunctions and alternations (see end of Chapter 22). '$\forall x\ \exists y\ Fxy$' becomes first:

$$\exists y\ Fay \,.\, \exists y\ Fby \,.\,.\,.\,.\, \exists y\ Fhy$$

and then:

$$Faa \vee Fab \vee \ldots \vee Fah \,.\, Fba \vee Fbb \vee \ldots \vee Fbh \,.\,.\,.\,. \\ Fha \vee Fhb \vee \ldots \vee Fhh.$$

On the other hand '$\exists y\ \forall x\ Fxy$' becomes first:

$$\forall x\ Fxa \vee \forall x\ Fxb \vee \ldots \vee \forall x\ Fxh$$

and then:

$$Faa \,.\, Fba \,.\,.\,.\,. Fha \,.\vee.\, Fab \,.\, Fbb \,.\,.\,.\,. Fhb \,.\vee.\,.\,.\,.\vee. \\ Fah \,.\, Fbh \,.\,.\,.\,. Fhh.$$

It was remarked in Chapter 22 that though in ordinary language the words 'something' and 'everything' masquerade as substantives, their behavior deviates from that of genuine substantives. Further examples of such deviation are provided by (7) and (8). For, (7) might be put into words as 'Everything is identical with something', and (8) as 'Something is identical with everything'. If 'everything' and 'something' really be-

haved like names, we should expect these two statements to be equivalent—and in fact we should expect both to be false. But actually, as seen, (7) is true and (8) false. Further, if 'nothing' and 'everything' were genuine names we should certainly expect 'Nothing is identical with everything' to be false; actually, however, this statement simply denies (8) and hence is true. Also we might expect 'Everything is identical with everything' to be equivalent to the truth 'Everything is identical with itself', whereas actually it expresses the falsehood:

(9) $\forall x \forall y$(x is identical with y).

One reason why quantificational analysis aids clear thinking is simply that the spurious substantives 'something', 'everything', and 'nothing' (and their variants 'somebody', 'everybody', 'nobody') give way to a less deceptive idiom.

The combination '$\forall x\ \forall y$' in (9) is not to be thought of as somehow a double quantifier; '$\forall x\ \forall y\ Fxy$' is simply a quantification of '$\forall y\ Fxy$' as a whole Whereas (1)–(8) show existential quantifications within universal ones and vice versa, (9) shows universal within universal.

Part III brings no new logical symbols for use in sentences. Already toward the end of Chapter 23 we had an example about Sadie's shoplifting that was the equal of any sentence formulable in Part III. Our new gains are not in sentences but in schemata. The schemata now before us are the *quantificational schemata* generally, not just the monadic ones. They comprise '*p*', '*q*', '*Fx*', '*Fy*', '*Gx*', '*Hxy*', '*Hyx*', '*Hxx*', '*Fxyz*', etc., and everything thence constructible by truth-function notations and quantifiers. The polyadic ingredients—'*Hxy*' and its suite—are what are new. (It is customary not to use the same term letter with different numbers of variables—thus '*Fx*' and '*Fxyz*'—within one schema or within the schemata of one problem. Conventions could be devised to accommodate such use.)

Though the enrichment touches only the schemata, it is a crucial one. As remarked, it enables us to establish new implications between sentences. The broadened domain of schemata sustains a broadened concept of validity, and hence of implication, which so exceeds the bounds of Part II as to resist any comparable treatment. For validity of quantificational schemata in general it is impossible to devise a decision procedure. (See Chapter 34.) The same applies of course to consistency, implication, and equivalence.

The definition of validity is as before: truth under all interpretations in

all nonempty universes. The definitions of consistency, implication, and equivalence follow suit. Interpretation of a term letter in the monadic case means settling what things in the universe the letter is to be true of. In polyadic cases it means settling what pairs of things in the universe, or what triples, etc., it is to be true of.

The impossibility of a decision procedure for validity will not deter us from developing procedures for proving validity. As noted in Chapter 13, the difference is that a decision procedure assures an affirmative or negative answer every time, while a proof procedure assures at best an eventual affirmative answer where an affirmative answer is in order. Proof procedures for general quantification theory will occupy most of Part III.

HISTORICAL NOTE: An algebra of relations was propounded by DeMorgan in 1864 and much improved in 1870 by Peirce. It is related to polyadic quantification theory much as the Boolean algebra of classes is related to monadic quantification theory, though, as Korselt showed in about 1914,[1] its scope is not quite so broad. By 1879, thanks to Frege, full quantification theory was at hand.

EXERCISES

1. Supposing the universe limited to $a, b, \ldots, h$, expand the quantifications into alternation and conjunction in each of the following examples:

$$\forall x\ \forall y\ Fxy, \quad \exists x\ \exists y\ Fxy, \quad \forall y\ \forall x\ Fxy, \quad \exists y\ \exists x\ Fxy.$$

   Do each in two stages.

2. Rewrite these with help of quantification:

   Every solid is soluble in some liquid or other,
   There is a liquid in which every solid is soluble.

3. Rewrite this example (DeMorgan's) with help of quantification:

   If all horses are animals then all heads of horses are heads of animals.

[1] See van Heijenoort, pp. 229, 233.

4. Express, with help of quantification, the likeliest interpretation of the statement:

She had a ring on every finger.

5. Where 'F' means 'harms', and the universe is mankind, put these unambiguously and idiomatically into words:

$$\exists x\ \forall y\ Fyx,$$
$$\forall x(\exists y\ Fxy \rightarrow Fxx),$$
$$\forall x[\forall y(Fyx \rightarrow Fxy) \rightarrow Fxx].$$

6. Supposing the universe to comprise just the points on an endless line, judge each of these statements as to truth value and explain your reasoning.

$$\exists x \forall y \exists z(x \text{ is between } y \text{ and } z),$$
$$\forall y \exists x \exists z(x \text{ is between } y \text{ and } z),$$
$$\exists x \exists z \forall y(x \text{ is between } y \text{ and } z).$$

7. Express each of these statements with help of quantification and indicate its truth value.

Nothing is identical with nothing,
Something is identical with something,
Everything is identical with nothing,
Nothing is identical with anything.

28
SUBSTITUTION EXTENDED

In Chapter 26 we saw that the derivation of sentences by substitution in monadic quantificational schemata is a rather complex and subtle affair, and we saw that it can be smoothly managed through substitution of abstracts for the term letters. Now that we have moved to polyadic schemata, worse complexity is to be expected. How should sentences

represented by '*Fxy*', '*Fyx*', and '*Fxx*' be related to one another? We must be able to tell e.g. that the sentence:

(1) $\exists x(y \text{ amuses } x \text{ more than } y \text{ amuses } y)$
$\vee\ \exists x(x \text{ amuses } y \text{ more than } y \text{ amuses } x)$
$.\!\rightarrow\ \exists x(x \text{ amuses } x \text{ more than } y \text{ amuses } x)$

is a correct result of substitution in the schema:

(2) $\exists x\, Fxy \vee \exists x\, Fyx .\!\rightarrow \exists x\, Fxx.$

Happily the solution is much the same as before. The appropriate auxiliary is again the term abstract, but polyadic: thus '$\{xy: \ldots x \ldots y \ldots\}$', '$\{xyz: \ldots x \ldots y \ldots z \ldots\}$', etc. To get (1), the abstract to substitute for '*F*' in (2) is:

$$\{zw: w \text{ amuses } z \text{ more than } y \text{ amuses } w\}.$$

We have only to put it for '*F*' in all three places and then proceed with concretion in obvious fashion: in the first of the three places the bound variables '*z*' and '*w*' of abstraction give way to '*x*' and '*y*', in the second to '*y*' and '*x*', and in the third to '*x*' and '*x*'.

The two restrictions in Chapter 26 continue to apply. Inspection confirms that the present example is in order on that score.

In the monadic case, term abstracts are a regimentation of ordinary relative clauses. In the polyadic case they are analogous; they are a polyadic extension of the relative clause. Our ordinary language lacks polyadic relative clauses, and one can see in part why. Relative clauses have most of their utility in contexts equivalent to quantification: contexts in which they are governed in effect by the functor '$\exists$' or '$\forall$'. But a polyadic relative clause offers no advantage when governed by '$\exists$' or '$\forall$', because the same work can be accomplished with nested monadic relative clauses subject separately to '$\exists$' or '$\forall$'. For instance:

$$\exists\{xy: x \text{ saved } y \text{ from drowning}\}$$

amounts in monadic clauses to 'There is someone who saved someone who was drowning', or:

$$\exists\{x: \exists\{y: x \text{ saved } y \text{ from drowning}\}\}.$$

In other words, the combination '$\exists\{xy$:' amounts merely to the iterated quantifiers '$\exists x\ \exists y$'; and similarly for '$\forall\{xy$:'. But if polyadic relative clauses offer little to ordinary discourse, they are a godsend to the theory of substitution.

Here, as in monadic term abstraction, I have drawn on the notation of set theory. The notation '$\{xy$: . . . x . . . y . . .$\}$' is standard for relations, construed as classes of ordered pairs. Remarks parallel to those at the end of Chapter 21 apply again here.

It was remarked at the end of Chapter 7 that validity may be ascribed not only to truth-functional schemata but also, by extension, to the sentences whose forms those schemata depict; but that it is well then to add the qualifier 'truth-functional'. Correspondingly a sentence obtainable by substitution in a valid quantificational schema is *quantificationally* valid. Such a sentence is true, or true for all values of its free variables. But it may or may not be truth-functionally valid; its truth may depend solely on its truth-functional structure, or it may depend partly on how the quantifiers are arranged.

We may also note an intermediate grade, *monadic* validity. A sentence is quantificationally valid if it can be got by substitution in a valid quantificational schema; it is monadically valid, more particularly, if it can be got by substitution in a valid quantificational schema which is monadic; and it is truth-functionally valid if it can be got by substitution in a valid truth-functional schema.

Similar remarks apply to implication on the part of sentences: it can be quantificational, and more particularly monadic, and still more particularly truth-functional. Similarly for inconsistency.

Substitution in schemata also takes another direction and serves another purpose: substitution of schemata in schemata to produce more schemata. A notable feature of such substitution was remarked upon at the truth-functional level in Chapters 6 and 13: substitution in valid schemata gives valid schemata.

Just now we noted that a quantificationally valid sentence might or might not be, more particularly, monadically valid or even truth-functionally valid. Now similar distinctions apply to quantificational schemata themselves. A valid polyadic schema may be *monadically valid,* in the sense of issuing by substitution from a valid monadic one. An example is '$\forall x\ Gxy \rightarrow Gyy$', which issues from '$\forall x\ Fx \rightarrow Fy$'. And of course a quantificational schema may be truth-functionally valid.

When we were getting sentences by substitution, what we substituted were sentences for sentence letters and term abstracts for term letters. To

get schemata by substitution, what we substitute are quantificational schemata for sentence letters and again term abstracts for term letters; but the term abstracts are now schematic—e.g. '$\{w\colon Gw \vee \exists z\, Hzw\}$'. Substitution of this for 'F' in '$\forall x\, Fx \rightarrow Fy$' proceeds in the way and under the restrictions already noted, and yields:

(3) $$\forall x(Gx \vee \exists z\, Hzx) \rightarrow .\, Gy \vee \exists z\, Hzy.$$

Another example: substitution of '$\{w\colon Gw \vee -Hw\}$' for 'F' and 'Gy' for 'p' in:

(4) $$\forall x(Fx \rightarrow p) \leftrightarrow .\, \exists x\, Fx \rightarrow p$$

yields:

(5) $$\forall x(Gx \vee -Hx \mathrel{.\!\!\rightarrow} Gy) \leftrightarrow .\, \exists x(Gx \vee -Hx) \rightarrow Gy.$$

The utility of substitution here, as in Part I, is as a means of generating valid schemata from valid schemata. E.g., since (3) and (5) were got by substitution in schemata which were seen in Chapter 25 to be valid, we conclude that (3) and (5) are valid.

Substitution can be depended upon to transmit validity for essentially the reasons already noted in Chapter 6. But it will be well to review the matter now in the new setting. To begin with let us see why it is that the above substitution in (4) yields a valid result. Validity of the result (5) means truth under all interpretations of 'G' and 'H' and the free variable 'y', within any nonempty universe. Suppose the universe fixed, then, and consider any particular choice $\mathfrak{I}$ of such interpretations; what we want to see is that (5) comes out true under $\mathfrak{I}$. To see this we derive from $\mathfrak{I}$ the following interpretations for the schematic letters of (4): we interpret 'F' as true of just the things that '$\{w\colon Gw \vee -Hw\}$' comes to be true of under $\mathfrak{I}$, and we interpret 'p' as having the truth value which 'Gy' comes to have under $\mathfrak{I}$. Being valid, (4) must come out true under these interpretations; hence (5), which simply repeats (4) under these interpretations, comes out true too.

More generally, suppose a sentence schema S' obtained by substitution in S. Each free variable and each schematic letter in S has a correspondent among the materials of S'; this correspondent is in each case either the same letter over again, or else a substituted sentence schema or term letter or abstract. Now given any choice $\mathfrak{I}$ of interpretations for the free vari-

ables and schematic letters of S', let us adopt as interpretation of each free variable or schematic letter of S the interpretation that has already accrued to its correspondent through $\mathfrak{I}$. S, so interpreted, matches S' as interpreted by $\mathfrak{I}$. Since this works for each choice of $\mathfrak{I}$, we see that S' is valid (or true for all interpretations) if S is.

The function of the two restrictions on substitution in Chapter 26 is to assure that the correspondents just now spoken of really correspond. Let us now have some examples showing how substitution can fail to transmit validity when the restrictions are violated.

Substitution of '$\{z\colon \exists y\ Gzy\}$' for '$F$' in '$\forall x\ Fx \rightarrow Fy$', in violation of the first restriction, would yield:

(6) $\forall x\ \exists y\ Gxy \rightarrow \exists y\ Gyy.$ (invalid)

That this is not valid, despite the validity of '$\forall x\ Fx \rightarrow Fy$', is seen by confining the universe to numbers and interpreting 'G' as 'is less than'; thereupon the antecedent of (6) becomes true ('for every number there is a greater') and the consequent false.

Substitution of '$\{z\colon Gxz\}$' for 'F' in '$\forall x\ Fx \rightarrow Fy$', in violation of the second restriction, would yield:

(7) $\forall x\ Gxx \rightarrow Gxy.$ (invalid)

That this is not valid may be seen by taking 'G' as 'identical with'; then (7) says 'If everything is identical with itself then x is identical with y', and this is clearly not true for every choice of x and y. Or, to restate this refutation in more explicit relation to the definition of validity: when we adopt a universe of two or more objects, and take one of these objects as interpretation of the free 'x' of (7) and a different one as interpretation of 'y', and interpret 'G' as 'is identical with', thereupon (7) becomes false.

In (3), the expressions '$Gx \vee \exists z\ Hzx$' and '$Gy \vee \exists z\ Hzy$' which supplanted the 'Fx' and 'Fy' of '$\forall x\ Fx \rightarrow Fy$' are symmetrical in '$x$' and '$y$': the one expression has 'x' where, and only where, the other has 'y'. In the invalid substitution which led to (7), on the other hand, the expressions 'Gxx' and 'Gxy' which supplanted 'Fx' and 'Fy' fail to show this symmetry; 'Gxy' does not have 'y' everywhere that 'Gxx' has 'x'. The reader must be warned that this asymmetry has nothing to do with the invalidity of (7). It is unnecessary, in general, for the expression supplanting 'Fx' to have 'x' where and only where the expression

supplanting 'Fy' has 'y'. It is quite proper, e.g., to substitute '$\{z\colon Gzy\}$' for 'F' in '$\forall x\, Fx \rightarrow Fy$' and infer the validity of:

(8) $$\forall x\, Gxy \rightarrow Gyy$$

(e.g., 'If everyone hates y then y hates himself'). Despite the asymmetry of 'Gxy' and 'Gyy' with respect to 'x' and 'y', (8) is a genuine special case of '$\forall x\, Fx \rightarrow Fy$', as a verbal comparison immediately reflects: 'If everything is an F then y is an F'; 'If everything is a G of y then y is a G of y'; 'If everyone is a Herbert-hater then Herbert is a Herbert-hater'.

To be assured of the correctness of a substitution, we need look only to these points: we must be able, on demand, to specify the actual sentence schema or term letter or n-adic abstract that is substituted for the sentence letter or n-adically occurring term letter; we must be sure that it has been introduced at each occurrence of the letter; we must be sure that at each point of introducing the abstract the particular variables there appended to the term letter have been put for the variables of abstraction; and finally we must be sure that the substitution has not led to new capturing of variables by quantifiers, contrary to the two restrictions.

Let us now shift from '$\forall x\, Fx \rightarrow Fy$' to another obviously valid schema, '$Fy \rightarrow \exists x\, Fx$'. From this we may proceed to:

(9) $$Gyy \rightarrow \exists x\, Gxy$$

by the legitimate substitution of '$\{z\colon Gzy\}$' for 'F'; but it would be illegitimate to substitute '$\{z\colon Gxz\}$' for 'F' and thus proceed to:

(10) $$Gxy \rightarrow \exists x\, Gxx.$$ (invalid)

An example of (9) is 'If Herbert hates himself then someone hates Herbert', which is quite unexceptionable; but an example of (10) is 'If Amos is uncle of Herbert then someone is uncle of himself'. Note that though the same principles of substitution are operative here as before, the pair (9) and (10) is rather opposite in appearance to (8) and (7). The valid (8) had unlike variables in the antecedent, but in the valid (9) the opposite is the case.

As far as substitution in our particular examples '$\forall x\, Fx \rightarrow Fy$' and '$Fy \rightarrow \exists x\, Fx$' is concerned, note that the net effect of the two restrictions is just this: the schemata that come to supplant 'Fy' and 'Fx' must be

alike except that the one has free 'y' wherever the other has free 'x'. Both may have additional free occurrences of 'y', as seen in (8) and (9). Thus think of '$\ldots y \ldots y \ldots$' and '$\ldots x \ldots y \ldots$' as the two schemata; we can put them for 'Fy' and 'Fx' in '$\forall x\, Fx \rightarrow Fy$' and '$Fy \rightarrow \exists x\, Fx$' by substituting '$\{z\colon \ldots z \ldots y \ldots\}$' for '$F$'. So, insofar as we are concerned merely with substitution in '$\forall x\, Fx \rightarrow Fy$' and '$Fy \rightarrow \exists x\, Fx$', we may omit all thought of abstracts, instead directly supplanting 'Fx' as a whole by any sentence schema S_x containing free 'x', and 'Fy' by a schema S_y which is like S_x except for having free 'y' in place of all free 'x'.

Such, then, is the notational relation between sentence schemata that are suited to the respective roles of '$\forall x\, Fx$' and 'Fy' in '$\forall x\, Fx \rightarrow Fy$', or of '$\exists x\, Fx$' and '$Fy$' in '$Fy \rightarrow \exists x\, Fx$'. It has a name: one schema is called an *instance* of the other. 'Fy' is an instance of '$\forall x\, Fx$' and '$\exists x\, Fx$'. The terminology applies also, of course, where letters other than 'x' and 'y' are used; thus 'Fw' is an instance of '$\forall y\, Fy$' and '$\exists y\, Fy$'. Furthermore, 'Gww' is an instance of '$\forall y\, Gwy$', of '$\exists y\, Gwy$', of '$\forall y\, Gyy$', of '$\exists y\, Gyy$', of '$\forall y\, Gyw$', and of '$\exists y\, Gyw$'. On the other hand 'Gwy' is not an instance of '$\forall y\, Gyy$' or '$\exists y\, Gyy$'.

The general description is this: an instance of a quantification exactly matches the old open schema that followed the quantifier, except that it may show a different variable in place of the recurrences of the variable of that quantifier. If it does show a different variable, it must show it in all the places (at least) where the old variable had been free in the old open schema that followed the quantifier. Moreover, it must show it *free* in those places. The reader will recognize in these requirements the effects of the restrictions on substitution for term letters.

A universal quantification implies each of its instances, and an existential quantification is implied by each of its instances. Such are the implications that come of substitution in the valid schemata '$\forall x\, Fx \rightarrow Fy$' and '$Fy \rightarrow \exists x\, Fx$'.

When we substitute for 'F' in the valid closed schemata:

(11) $\forall y(\forall x\, Fx \rightarrow Fy)$, (13) $\exists y(Fy \rightarrow \forall x\, Fx)$,
(12) $\forall y(Fy \rightarrow \exists x\, Fx)$, (14) $\exists y(\exists x\, Fx \rightarrow Fy)$

(of Chapter 25, especially Exercise 1), the effect of the two restrictions is more stringent: the schemata that supplant 'Fy' and 'Fx' here must be alike except that the one has free occurrences of 'y' where *and only* where the other has free 'x'. For, the second restriction requires that the abstract

substituted for 'F' be devoid of free 'y', in view of the initial quantifiers in (11)–(14).

Thus, whereas it was allowable to substitute '$\{z\colon Gzy\}$' for 'F' in '$\forall x\, Fx \to Fy$' and '$Fy \to \exists x\, Fx$' so as to obtain (8) and (9), it is forbidden to make the same substitution in (11)–(14) so as to obtain:

(15) $\forall y(\forall x\, Gxy \to Gyy)$, (17) $\exists y(Gyy \to \forall x\, Gxy)$, (invalid)
(16) $\forall y Gyy \to \exists x\, Gxy)$, (18) $\exists y(\exists x\, Gxy \to Gyy)$. (invalid)

(15) and (16) happen to be valid anyway, because they are the universal closures of the valid open schemata (8) and (9). (17) is not valid, as may be seen by adopting a universe of two or more objects and interpreting 'G' as 'is identical with'. 'Gyy thereupon becomes true and '$\forall x\, Gxy$' false for every object y, and thus (17) comes out false. That (18) is not valid may be seen by interpreting 'G' as 'is distinct from'.

HISTORICAL NOTE: I know of no full and correct formulation of substitution for polyadic term letters prior to that of 1934 by Hilbert and Bernays. The auxiliary role played by abstracts in my present method was played in theirs by schemata called *Nennformen*. In *Elementary Logic* I used auxiliary formulas of another style for that purpose. They contained circled numerals and were called *stencils*, or, in the second edition of that book and three editions of this one, *predicates*. It is a pleasure now to invoke the versatile abstract and banish those intrusions.

The two restrictions on substitution date from *Elementary Logic*, 1941; Hilbert and Ackermann had met those needs by the more drastic expedient of using different parts of the alphabet for their bound and free variables.

EXERCISES

1. Decide which of the following abstracts may be substituted for 'F' in '$\exists x\, Fxy$' compatibly with the restrictions on substitution.

$\{zw\colon w$ praised w to $z\}$, $\{zw\colon w$ praised y to $z\}$,
$\{zw\colon w$ praised z to $z\}$, $\{zw\colon \exists y(w$ praised y to $z)\}$,
$\{zw\colon w$ praised x to $z\}$, $\{zw\colon \exists u(w$ praised u to $z)\}$.

Put the results of those legitimate substitutions into words, supposing the universe limited to mankind.

2. List all the schemata whose validity can be shown by legitimate substitution of one or another of:

$$\{z\colon Gxz \vee Gzy\}, \qquad \{z\colon Gyz \vee Gzy\},$$
$$\{z\colon Gxz \vee Gzx\}, \qquad \{z\colon \forall y(Gyz \vee Gzy)\}$$

for 'F' in '$\forall x\, Fx \rightarrow Fy$' or '$Fy \rightarrow \exists x\, Fx$' or (4) or (13) or (14). Taking the universe as the members of the council and interpreting 'G' as 'denounced' and 'p' as 'steps must be taken', put the results into ordinary language.

3. Determine which of the following schemata are legitimately obtainable from '$\forall x\, Fx \rightarrow Fy$' or '$Fy \rightarrow \exists x\, Fx$' or (4) or (13) or (14) by substitution. Identify the substituted abstract in each case.

$\forall x\, Fxx \rightarrow Fyy$, $\qquad Gyx \rightarrow \exists x\, Gxy$,
$\exists x(Gxx \rightarrow \exists y\, Gyy)$, $\qquad \exists y(Gy \,.\, Hz \,.\!\rightarrow\!.\, \forall x\, Gx \,.\, Hz)$,
$\forall x(Gyx \rightarrow Hxz) \rightarrow.\, Gyy \rightarrow Hyz$, $\qquad Gxy \,.\, Hyz \,.\!\rightarrow\, \exists x(Gxx \,.\, Hxz)$,
$Gy \,.\, \forall z(Gy \rightarrow Gz) \,.\!\rightarrow\, \exists x[Gx \,.\, \forall z(Gx \rightarrow Gz)]$,
$\forall x[\forall y(Gxy \rightarrow Gyy) \rightarrow Gyy] \leftrightarrow.\, \exists x \forall y(Gxy \rightarrow Gyy) \rightarrow Gyy$.

29
PURE EXISTENTIALS

In general quantification theory as elsewhere, implication is validity of the conditional. To justify an inference, then, such as the one in Chapter 27 about the philosopher, what we need to do is prove the validity of a conditional. In this example the appropriate conditional is:

(1) $$\exists y[Fy \,.\, \forall x(Fx \rightarrow Gxy)] \rightarrow \exists x(Fx \,.\, Gxx)$$

(see (4) in Chapter 27). Converting it to prenex form, we get perhaps:

$$\forall y\, \exists x\, \exists z(Fy \,.\, Fx \rightarrow Gxy \,.\!\rightarrow\!.\, Fz \,.\, Gzz).$$

Other prenex equivalents could be got, but this will do. In fact, since validity is our concern, we may as well drop the '$\forall y$'; for an initial universal quantifier is indifferent to validity. What we want to show, then, is the validity of:

(2) $$\exists x \exists z(Fy \,.\, Fx \rightarrow Gxy \,.\!\rightarrow\!.\, Fz \,.\, Gzz).$$

Now try this experiment: substitute the free 'y' for the existential variables, dropping the quantifiers. The result:

(3) $$Fy \,.\, Fy \rightarrow Gyy \,.\!\rightarrow\!.\, Fy \,.\, Gyy$$

is truth-functionally valid. This outcome, properly viewed, establishes the validity of (2) and therewith of (1). For (3) is an instance of:

$$\exists z(Fy \,.\, Fy \rightarrow Gyy \,.\!\rightarrow\!.\, Fz \,.\, Gzz)$$

and accordingly implies it, which in turn implies (2).

The *matrix* of a prenex schema is what comes after the quantifiers. A *pure existential* is a prenex schema with none but existential quantifiers. Now what the example above teaches us is that if a pure existential has just one free variable, and substitution of that variable for the existential variables turns the matrix truth-functionally valid, then the pure existential was valid.

We can claim more: the pure existential is valid *only* if that substitution turns the matrix truth-functionally valid. The reason may best be seen through another example. Let us modify (2) just enough to spoil it.

(2′) $$\exists x \exists z(Fy \,.\, Fx \vee Gxy \,.\!\rightarrow\!.\, Fz \,.\, Gzz).$$

Substitution of 'y' for 'x' and 'z' in the matrix of (2′) gives:

(3′) $$Fy \,.\, Fy \vee Gyy \,.\!\rightarrow\!.\, Fy \,.\, Gyy,$$

which, rather than being truth-functionally valid, is falsifiable by assigning $\top$ to 'Fy' and $\bot$ to 'Gyy'. But (2′) itself boils down to (3′) in a universe of just a single object y. So (2′) comes out false in that universe if we interpret 'F' and 'G' so that Fy and $-Gyy$. So (2′) is not valid: not true under all interpretations of 'F' and 'G' in all nonempty universes.

We have arrived at a decision procedure for pure existentials in one free variable. The schema is valid if and only if its matrix becomes truth-functionally valid on substitution of the free variable for the existential variables.

The test can be extended to pure existentials in more than one free variable. In such a case there are a multiplicity of ways of substituting the free variables for the existential ones; but then we take the alternation of all such results and check it for truth-functional validity. Example:

(4) $$\exists z(Fxz \rightarrow Fzy).$$

There are two ways of substituting the free variables for the existential variable. They give '$Fxx \rightarrow Fxy$' and '$Fxy \rightarrow Fyy$'. The alternation:

(5) $$Fxx \rightarrow Fxy \mathbin{.\vee.} Fxy \rightarrow Fyy$$

is truth-functionally valid. Each half of (5) implies (4); so any interpretation that makes either half true makes (4) true. In short, (5) implies (4). So (4) is valid. It is strange but true, logically true, for any two people John and Mary, that there is someone who, if admired by John, admires Mary.

To see in general that the validity of the alternation is not only sufficient for the validity of the pure existential but also necessary, let us again reason from an example. Let us modify (4) just enough to spoil it.

(4′) $$\exists z(Fxz \rightarrow Fyz).$$

The alternation:

(5′) $$Fxx \rightarrow Fyx \mathbin{.\vee.} Fxy \rightarrow Fyy$$

is falsifiable by assigning T to 'Fxx' and 'Fxy' and $\perp$ to 'Fyx' and 'Fyy'. But (4′) itself boils down to (5′) if we take a universe of just x and y. So (4′) comes out false in that universe under such an interpretation of 'F'.

At this point the form of argument is evident for this general theorem: *A pure existential is valid if and only if we get a truth-functionally valid schema by taking the alternation of the results of substituting the free variables for the existential ones in the matrix.*

In hopes of checking validity along these lines, we do well, when converting a schema to prenex form, to choose our steps so as to give priority to universal quantifiers where possible. If we can get them all in

front, we can drop them and test the pure existential. If in converting (1) we had brought its final '$\exists x$' out first, this test would have been denied us.

Two fringe cases need notice. What if on conversion to prenex form all the quantifiers become universal, and thus are dropped? Then of course we test the schema as it stands for truth-functional validity, there being no quantifiers. And what if we end up rather with just existential quantifiers and no free variables? Then we just put the arbitrary letter 'x' for the variables in the matrix and test the resulting one-variable schema for truth-functional validity. E.g., the validity of:

$$\exists x \exists y(Fxy \rightarrow Fyx)$$

comes down to that of '$Fxx \rightarrow Fxx$'. The criterion is sufficient, since the one-variable schema is an instance (of an instance . . .) of the existential schema and so implies it. And it is necessary, since, in the universe of a single object x, the existential schema boils down to the one-variable schema.

Valid polyadic schemata may or may not be monadically valid (see Chapter 28). The validity of (1) and (4) is an irreducibly polyadic matter. The same is true of:

(6) $$\exists x \, \forall y \, Fxy \rightarrow \forall y \, \exists x \, Fxy,$$

which tends to be the first genuinely polyadic example that one thinks of. Let us try our new method on it. Converting it to prenex form with priority on universal quantifiers, we get:

$$\forall x \forall y \exists z \exists w(Fxz \rightarrow Fwy)$$

and so proceed to test this:

(7) $$\exists z \exists w(Fxz \rightarrow Fwy).$$

One way of substituting the free variables of (7) for 'z' and 'w' in '$Fxz \rightarrow Fwy$' gives '$Fxy \rightarrow Fxy$'; and there are three others. The test of validity of (7), and hence of (6), would consist in forming the alternation of '$Fxy \rightarrow Fxy$' with its three companions and examining the result for truth-functional validity. But of course there is no need of all this, since '$Fxy \rightarrow Fxy$' is valid by itself; the alternation has to be valid.

It happens conveniently often, in testing pure existentials, that a single such result of substitution turns out thus to be valid by itself. Thus take the example in Chapter 27 about drawing circles. Here the conditional to test for validity is:

(8) $\forall x(Fx \rightarrow Gx) \rightarrow \forall y[\exists x(Fx \,.\, Hyx) \rightarrow \exists x(Gx \,.\, Hyx)]$

(see Chapter 27, (1)). Converting it to prenex form with priority on universal quantifiers and then dropping the initial universal quantifiers, we get the pure existential:

(9) $\exists z\, \exists w(Fz \rightarrow Gz \mathrel{.\!\rightarrow\!:} Fx \,.\, Hyx \mathrel{.\!\rightarrow\!.} Gw \,.\, Hyw).$

Now there are four ways of substituting 'x' and/or 'y' for 'z' and 'w' in the matrix of (9). One way gives:

(10) $Fx \rightarrow Gx \mathrel{.\!\rightarrow\!:} Fx \,.\, Hyx \mathrel{.\!\rightarrow\!.} Gx \,.\, Hyx,$

and there are three others. The test of validity of (9), and hence of (8), would consist in forming the alternation of (10) with the other three and testing the long result for truth-functional validity. But again we do better to look before leaping; (10) itself is truth-functionally valid, so of course the alternation will be so likewise.

The reader will have the same experience again if he tries the second example in Chapter 27, the one about the paintings; again a single result of substitution turns out valid. Sometimes, however, we have to press on to the alternation of two or more such results of substitution. This was true already of the simple example (4); neither half of (5) is valid by itself.

So we have a decision procedure for validity of pure existentials. In view of the warning that there is no general decision procedure for validity of quantificational schemata, then, we may be sure that universal quantifiers cannot always be maneuvered into initial position. Here is a simple valid schema that resists such treatment:

(11) $\exists y \forall z(Fxz \rightarrow Fyz).$

To prove the validity of such schemata we shall have to resort to more general methods.

But it is a good policy, when trying to prove the validity of a schema, to try first for a pure existential. This often succeeds, and when it does there are two advantages. One is that a direct test is available, assuring a

conclusively negative answer or a proof. The other is that the process tends to be relatively swift, since it so often happens, as in all three of the examples that we have been examining from Chapter 27, that a single substitution of universal variables for the existential ones settles matters.

Incidentally we now have another general decision procedure for validity of monadic schemata, alternative to the methods in Part II. For we saw in Chapter 24 how any monadic schema can be purified; and we saw that, once purified, it will have no stacked quantifiers. But the rules of passage enable us to bring quantifiers into prenex position in any order we like if none of them are stacked. Thus any monadic schema can be tested for validity by the method of pure existentials.

By way of opening a field for further examples involving polyadic schemata, let us acquaint ourselves with *symmetry, transitivity, reflexivity,* and related concepts—these being worth noting also in their own right. A dyadic term is called symmetrical, asymmetrical, transitive, intransitive, totally reflexive, reflexive, or irreflexive according as it fulfills:

$\forall x \forall y(Fxy \rightarrow Fyx)$	(symmetry)
$\forall x \forall y(Fxy \rightarrow -Fyx)$	(asymmetry)
$\forall x \forall y \forall z(Fxy \,.\, Fyz \,.\!\rightarrow\!. Fxz)$	(transitivity)
$\forall x \forall y \forall z(Fxy \,.\, Fyz \,.\!\rightarrow\!. -Fxz)$	(intransitivity)
$\forall x\, Fxx$	(total reflexivity)
$\forall x \forall y(Fxy \,.\!\rightarrow\!.\, Fxx \,.\, Fyy)$	(reflexivity)
$\forall x\, -Fxx$	(irreflexivity)

The dyadic term 'compatriot' is symmetrical, in that if x is a compatriot of y then y is a compatriot of x. It is also transitive, if we disallow multiple nationality; for then if x is a compatriot of y and y of z, x will also be a compatriot of z. It is also reflexive, if we consider a person a compatriot of himself—as indeed we must if 'compatriot of' means 'having same nationality as'. But it is not totally reflexive, if we think of our universe as containing any things devoid of nationality. Examples of total reflexivity are rare and trivial; 'identical' and 'coexistent' are two such.

The dyadic term 'north' is again transitive, but it is asymmetrical and irreflexive; 'x is north of y' excludes 'y is north of x', and nothing is north of itself. The dyadic term 'mother' is intransitive, asymmetrical, and irreflexive.

The dyadic term 'loves' lacks all seven properties. Where x loves y, y may or may not love x; thus 'loves' is neither symmetrical nor asymmetrical. Where x loves y and y loves z, x may or may not love z; thus 'loves' is

neither transitive nor intransitive. And, since some love themselves while others (even among those who love or are loved) do not love themselves, 'loves' is neither reflexive nor irreflexive.

The reader may wonder why, parallel to the distinction between reflexivity and total reflexivity, a distinction is not drawn between "irreflexivity" in the sense of:

$$\forall x \forall y (Fxy \rightarrow . \sim Fxx . \sim Fyy)$$

and "total irreflexivity" in the sense of '$\forall x \sim Fxx$'. The reason is that this latter distinction is illusory; the two schemata are equivalent. To show this we establish the validity of two conditionals, as follows:

(12) $\forall x \forall y (Fxy \rightarrow . \sim Fxx . \sim Fyy) \rightarrow \forall x \sim Fxx,$

(13) $\forall x \sim Fxx \rightarrow \forall x \forall y (Fxy \rightarrow . \sim Fxx . \sim Fyy).$

Turning them into prenex form and dropping initial universal quantifiers, we get:

(14) $\exists x \exists y (Fxy \rightarrow . \sim Fxx . \sim Fyy :\rightarrow \sim Fzz),$

(15) $\exists z (\sim Fzz \rightarrow : Fxy \rightarrow . \sim Fxx . \sim Fyy).$

The one is shown valid by this truth-functionally valid result of substitution:

$$Fzz \rightarrow . \sim Fzz . \sim Fzz :\rightarrow \sim Fzz$$

and the other by this truth-functionally valid alternation of two:

$$\sim Fxx \rightarrow : Fxy \rightarrow . \sim Fxx . \sim Fyy$$
$$. : \vee : . \sim Fyy \rightarrow : Fxy \rightarrow . \sim Fxx . \sim Fyy.$$

As another example it will be proved that symmetry and transitivity together imply reflexivity. This means proving the validity of the conditional:

(16) $\forall x \forall y (Fxy \rightarrow Fyx) . \forall x \forall y \forall z (Fxy . Fyz .\rightarrow Fxz)$
$$.\rightarrow \forall x \forall y (Fxy \rightarrow . Fxx . Fyy).$$

Turning it into prenex form and dropping initial universal quantifiers, we get:

(17) $\exists v \exists w \exists x \exists y \exists z (Fvw \rightarrow Fwv : Fxy \,.\, Fyz \,.\!\rightarrow Fxz$
$:\!\rightarrow\!: Ftu \rightarrow . Ftt \,.\, Fuu)$.

This is shown valid by the truth-functional validity of the alternation of two of the results of substituting 't' and 'u' for the existential variables.

(18) $Ftu \rightarrow Fut : Ftu \,.\, Fut \,.\!\rightarrow Ftt :\!\rightarrow\!: Ftu \rightarrow . Ftt \,.\, Fuu$
$.:\vee:. Ftu \rightarrow Fut : Fut \,.\, Ftu \,.\!\rightarrow Fuu :\!\rightarrow\!: Ftu \rightarrow . Ftt \,.\, Fuu$.

The reader is urged to do a painstaking truth-value analysis before agreeing to the truth-functional validity of this alternation.

HISTORICAL NOTE: The existence of a validity test to this purpose seems to have been first noted by Bernays and Schönfinkel, 1928. In the present method there are echoes of Herbrand that will become clearer in Chapter 36.

EXERCISES

1. Delineate all the steps of passage, relettering, and deletion that carry (1) into (2); (6) into (7); (8) into (9); (12) into (14); (13) into (15); (16) into (17).
2. Check the validity of (18).
3. Do the example in Chapter 27 about the paintings.
4. Repeat the exercises of Chapter 25 by the new method.
5. Prove that asymmetry implies irreflexivity.
6. Prove that intransitivity implies irreflexivity.
7. Prove that transitivity and irreflexivity together imply asymmetry.

30
THE MAIN METHOD

We turn now to a proof procedure that will be found to be complete: adequate to establishing the validity of any valid quantificational schema and hence also to establishing any implication and any inconsistency. In fact it will be oriented to inconsistency proofs. To prove a schema valid we shall prove its negation inconsistent; and to prove implication we shall prove that the one schema is inconsistent with the negation of the other.

In (11) of the preceding chapter we saw a candidate for a validity proof, inaccessible to the method of pure existentials. To prove it valid by the new method we negate it, assume the prenex form of the negation as premise, and proceed to generate instances from it as follows:

PREMISE: $\forall y\ \exists z\ -(Fxz \rightarrow Fyz)$
INSTANCES: $\exists z\ -(Fxz \rightarrow Fxz)$
$-(Fxu \rightarrow Fxu)$

Proofs of inconsistency in this system always end up with a truth-functionally inconsistent schema, as here, or a truth-functionally inconsistent combination of schemata; and this melancholy exhibit is meant to show that the premise was at fault and thus inconsistent. Each successive line is an instance of an earlier line.

We seem to see here the pattern traditionally known as *reductio ad absurdum:* disproof by derivation of a clear contradiction. But derivation how? Do quantifications imply their instances?

Universal quantifications do indeed. The first line of the proof above implies the second. The operation of *universal instantiation*—UI, as I shall call it—which leads from the first line to the second is implicative.

But *existential* instantiation, EI, which leads from the second line to the last, is not implicative. '$\exists x\ Fx$' does not imply 'Fy'. Consequently it is not evident that the derivation of an inconsistency by progressive instantiation, as in the proof above, suffices to show the premise inconsistent. It does suffice, but the point will take some proving.

Before proceeding to justify the method, let us get a firmer idea of what the method is. For one thing, EI is to be hedged thus: The *instantial variable,* which is substituted for the variable that was bound by the

dropped existential quantifier, must be *new*. More accurately: it must be free in no line prior to this instantiation. Thus consider how unfortunate it would be to argue as follows:

PREMISES: $\exists x\, Fx$
$\exists x\, -Fx$
INSTANCES: Fy
$-Fy$ (wrong)

Another point, this time one of liberality rather than of hedging, is that we can have more than one premise, as above. For a valid example, take the familiar implication that was proved in the preceding chapter by proving the validity of (6). To prove it by the present method we would assume as premises '$\exists x\ \forall y\ Fxy$' and the prenex form of the negation of '$\forall y\ \exists x\ Fxy$', and proceed to instantiate to a truth-functional inconsistency:

PREMISES: $\exists x\ \forall y\ Fxy$
$\exists y\ \forall x\ -Fxy$
INSTANCES: $\forall y\ Fzy$
$\forall x\ -Fxw$
Fzw
$-Fzw$

As a further simple example let us prove that '$\forall x\ \forall y\ Fxy$' implies '$\forall y\ Fyy$'.

PREMISES: $\forall x\ \forall y\ Fxy$
$\exists y\ -Fyy$
INSTANCES: $-Fzz$
$\forall y\ Fzy$
Fzz

The interesting point about this implication is that one is apt to say that it is obvious by direct instantiation, '$\forall y\ Fyy$' being simply an instance of '$\forall x\ \forall y\ Fxy$'. One would be wrong; one could as well argue that '$\forall x\ \exists y\ Fxy$' implies '$\exists y\ Fyy$', but it does not. See (6) of Chapter 28. '$\forall y\ Fyy$' and '$\exists y\ Fyy$' are not instances of '$\forall x\ \forall y\ Fxy$' and '$\forall x\ \exists y\ Fxy$'. The definition of instance at the end of Chapter 28 needs to

be carefully observed and the reasons behind it, namely the restrictions on substitution, fully appreciated.

Here, next, is a more substantial example. It will serve as a source of illustration when we get to justifying the method.

PREMISES:	(1)	$\exists u\, \forall y\, Fuy$	
	(2)	$\forall x\, \exists v\, \forall y - (Fvy \,.\, Fyx)$	
	(3)	$\forall x \forall y \exists w (Fxy \rightarrow: -Fxw \,.\, Fwy \,.\vee.\, Fyw \,.\, Fwx)$	
INSTANCES:	(4)	$\forall y\, Fzy$	[of (1)]
	(5)	$\exists v\, \forall y - (Fvy \,.\, Fyz)$	[of (2)]
	(6)	$\forall y - (Fty \,.\, Fyz)$	[of (5)]
	(7)	Fzt	[of (4)]
	(8)	$\forall y \exists w (Fzy \rightarrow: -Fzw \,.\, Fwy \,.\vee.\, Fyw \,.\, Fwz)$	[of (3)]
	(9)	$\exists w (Fzt \rightarrow: -Fzw \,.\, Fwt \,.\vee.\, Ftw \,.\, Fwz)$	[of (8)]
	(10)	$Fzt \rightarrow: -Fzs \,.\, Fst \,.\vee.\, Fts \,.\, Fsz$	[of (9)]
	(11)	Fzs [of (4)]	
	(12)	$-(Fts \,.\, Fsz)$ [of (6)]	

(7), (10), (11), and (12) are together inconsistent by truth-value analysis. We conclude that (1)–(3) are together inconsistent.

The proof procedure for inconsistency that I have now described will be called the *main method,* to distinguish it from alternatives in Chapters 29 and 35–38. Summed up, it is the following slight affair. Instantiation: apply UI and EI, always choosing new instantial variables for EI. Termination: a truth-functionally inconsistent assemblage of unquantified instances.

I have still to show that the method is sound—i.e., that it generates a truth-functional inconsistency only when the premises are inconsistent. In showing this I shall need some terminology. I shall use 'generic' as the correlative of 'instance'; thus a step of UI or EI derives an instance from its generic. I shall speak of U-instances and E-instances. Thus in the inconsistency proof (1)–(12) the E-instances are (4), (6), and (10), and their generics are (1), (5), and (9). By an EI-*conditional* I shall mean a conditional whose consequent and antecedent copy an E-instance and its generic. If for compactness I simply write the numerals themselves as abbreviations of the numbered lines, then the EI-conditionals of the inconsistency proof (1)–(12) are '1 → 4', '5 → 6', and '9 → 10'.

Now observe that each instance, in such a proof, is implied by previous lines plus EI-conditionals. For each U-instance is implied by its generic, and each E-instance is truth-functionally implied by its generic and its EI-conditional. (E.g., (4) is implied by (1) and '1 → 4'.) Ultimately,

therefore, all the instances are implied by the premises and the EI-conditionals. So, since the instances are inconsistent, the conjunction of the premises and the EI-conditionals is inconsistent. In our example this conjunction is:

$$1 \,.\, 2 \,.\, 3 \,.\, 1 \rightarrow 4 \,.\, 5 \rightarrow 6 \,.\, 9 \rightarrow 10.$$

As noted early in Chapter 25, it will remain inconsistent when an existential quantifier is prefixed. Let us prefix one containing the instantial variable of the last E-instance.

$$\exists s(1 \,.\, 2 \,.\, 3 \,.\, 1 \rightarrow 4 \,.\, 5 \rightarrow 6 \,.\, 9 \rightarrow 10).$$

Now we use the fact that this instantial variable, having been new (i.e., free in no prior line), is free nowhere in our conjunction except in the last EI-conditional, and only in its consequent at that. We can confine the quantifier to that part by the rules of passage (1) and (5) of Chapter 23.

$$1 \,.\, 2 \,.\, 3 \,.\, 1 \rightarrow 4 \,.\, 5 \rightarrow 6 \,.\, 9 \rightarrow \exists s\ 10.$$

But (9) and '$\exists s$ 10' are alike except for the alphabetical choice of the existential variable. So '$9 \rightarrow \exists s$ 10' is valid and can be dropped from the conjunction, leaving '$1 \,.\, 2 \,.\, 3 \,.\, 1 \rightarrow 4 \,.\, 5 \rightarrow 6$'. Then by a similar argument we get rid of '$5 \rightarrow 6$' and finally of '$1 \rightarrow 4$'. Thus the whole inconsistent conjunction has boiled down to the premises, which are therefore inconsistent, Q.E.D. In this argument I have leaned increasingly on the particular example, but still the general reasoning shines through.

We see thus that our main method is a sound one. It is a rather natural one, moreover, despite the austerity of the argument above. The pattern (1)–(12) could be verbalized, as a disproof of some conjunction of actual statements in lieu of (1)–(3), along the following line. According to (1), there is something that is F to everything. Very well, call it z. So we have (4). But (2) said that everything x is such that . . . . . Well then in particular z will be that way. So we have (5): that there is something such that . . . . . Call it t. Continuing thus, we get to the contradiction (7), (10)–(12). The premises (1)–(3) are thus disproved. One or two of them may be true, but not all three.

Premises must be put in readiness in two respects before the main method is applied to them. They must be converted to prenex form; this we know. The further point is that the letters used as bound variables in

each premise should be so chosen as not to match a free variable of any premise. E.g., consider '$\exists w\ -Fyw$' and '$\forall x\ \forall y\ Fxy$'. They are inconsistent, but the reader can try in vain to show them inconsistent by the main method as they stand; whereas the proof goes smoothly through if the bound 'y' is rewritten as 'z'.

PREMISES:	$\exists w\ -Fyw$
	$\forall x\ \forall z\ Fxz$
INSTANCES:	$-Fyu$
	$\forall z\ Fyz$
	Fyu

We shall see in Chapter 32 that the main method, for all its slightness, is complete. We shall see that for every inconsistent conjunction of schemata, prepared in the two respects just now indicated, it affords an inconsistency proof.

HISTORICAL NOTE: Reductio ad absurdum, also called *indirect proof,* was known to the ancients under the name of *apagoge*. Its advantage over direct proof of validity turns only on a small point of technical convenience which will be noted in Chapter 36. The first proof procedure for quantification theory—Frege's of 1879—proceeded rather in the axiomatic style of Chapter 13; see Chapter 37 below. This style continued to prevail through the thirties. Already in 1928 and 1930, however, Skolem and Herbrand presented proof procedures more akin to what I am calling the main method; see Chapters 35 and 36. The main method figured in a 1955 appendix to earlier editions of the present book.

EXERCISES

1. Find fault with the following purported proofs of inconsistency.

PREMISES:	$\exists x\ Fxy$	$-\exists x\ Fxx$	$\exists x\ Fxy$
	$\forall x\ -Fxx$	$\forall y\ Fyy$	$\exists x\ -Fxy$
INSTANCES:	Fyy	$-Fzz$	Fzy
	$-Fyy$	Fzz	$-Fzy$

Is any of these pairs of premises really inconsistent? Which? Prove it so.

2. Repeat Exercises 3–7 of Chapter 29 by the main method.

3. Describe the obstacle to showing '$\exists w\ -Fyw$' and '$\forall x\ \forall y\ Fxy$' inconsistent by the main method as they stand.

31
APPLICATION

When our proof procedure is to be brought to bear upon statements couched in ordinary language, the task of suitably paraphrasing the statements and isolating their relevant structure becomes just as important as the proof for which that preliminary task prepares the way (see Chapter 8).

In Chapter 14 we noted a considerable variety of ways in which the categorical forms **A**, **E**, **I**, and **O** may appear in ordinary language; and in Chapter 22 we saw how to put those forms over into quantificational notation. These observations provide the beginning of a guide to the translation of words into quantificational symbols. But we saw also, from examples such as 'A lady is present', 'A Scout is reverent', 'John cannot outrun any man on the team', and 'Tai always eats with chopsticks' (Chapter 14), that it is a mistake to trust to a pat checklist of idioms. The safer way of paraphrasing words into symbols is the harder way: by a sympathetic rethinking of the statement in context. If there are obvious ways of rectifying logically obscure phrases by rewording, it is well to do so before resorting to logical symbols at all.

A drastic departure from English is required in the matter of tense. The view to adopt is the Minkowskian one, which sees time as a fourth dimension on a par with the three dimensions of space. Quantifiers must be read as timeless. The values of 'x' may themselves be thing-events, four-dimensional denizens of space-time, and we can attribute dates and durations to them as we can attribute locations and lengths and breadths to them; but the quantifier itself attributes none of these things. '$\exists x$' says neither 'there was' nor 'there will be', but only, in a tenseless sense, 'there is'.

This four-dimensional view was needed, as everyone knows, to make sense of Einstein's relativity physics. But it also has long been a great

help in clarifying ordinary talk on more humdrum matters. When we say how many presidents or popes there have been, we state the size of a class whose members have never all coexisted. When we compare Napoleon to Caesar or trace David's descent from Abraham, we relate persons who never coexisted. Puzzles both conceptual and verbal are minimized by seeing spatial and temporal associations as logically alike. Much as Boston and Birmingham are 3000 miles apart, Caesar and Napoleon are 1800 years apart; and the 'are' here had better be read as tenseless.

From long habituation we are proof against silly fallacies of this sort: George V married Queen Mary, Queen Mary is a widow, therefore George V married a widow. Clearly, however, a logic designed to control this sort of thing explicitly would be needlessly elaborate. We do better to make do with a simpler logical machine and then, when we want to apply it, to paraphrase our sentences to fit it. Already in Chapter 14 we had a glimpse of the varied ways in which temporal references might fare under such paraphrase. Not, of course, that the ubiquity of tense in English will require us to make explicit reference to time in all the paraphrases. As often as not the temporal matter is superfluous, and foisted on us only by English usage. Still more often, in practice, we can even leave the tensed verbs themselves undisturbed, as long as there is no danger of equivocation within the space of the proof. (See Chapter 8.)

In paraphrasing more complex statements into quantificational form, a problem that obtrudes itself at every turn is that of determining the intended groupings. The cues to grouping which were noted at the truth-functional level in Chapter 4 continue to be useful here, but the most important single cue proves to be the additional one which was noted in connection with (11) of Chapter 22: *The scope of a quantifier must reach out far enough to take in any occurrence of a variable that is supposed to refer back to that quantifier.*

The technique of *paraphrasing inward* (Chapter 8), as a means of dividing the problem of interpretation into manageable parts and keeping the complexities of grouping under control, is as important here as at the truth-functional level; more important, indeed, in proportion to the increasing complexity of the statements concerned. After each step of paraphrasing, moreover, it is well to check the whole against the original statement to make sure that the intended idea is still being reproduced.

By way of a serious venture in paraphrasing, let us try putting the following premises and conclusion over into quantificational form preparatory to setting up a deduction.

PREMISES: The guard searched all who entered the building except those who were accompanied by members of the firm,
Some of Fiorecchio's men entered the building unaccompanied by anyone else,
The guard searched none of Fiorecchio's men.
CONCLUSION: Some of Fiorecchio's men were members of the firm.

The first premise says in effect:

> Every person that entered the building and was not searched by the guard was accompanied by some member(s) of the firm.

Setting about now to paraphrase this premise inward, we inspect it for its outermost structure, which obviously is '$\forall x(\ldots \rightarrow \ldots)$':

> $\forall x$(x is a person that entered the building and was not searched by the guard $\rightarrow x$ was accompanied by some members of the firm).

The virtue of thus paraphrasing inward a step at a time is that the unparaphrased internal segments can now be handled each as a small independent problem. The clause 'x was accompanied by some members of the firm', e.g., regardless of context, becomes:

> $\exists y$(x was accompanied by y . y was a member of the firm).

The other clause, 'x is a person that entered the building and was not searched by the guard', needs little more attention; we have merely to make it an explicit conjunction:

> x is a person that entered the building . x was not searched by the guard.

So the whole becomes:

> $\forall x$[x is a person that entered the building . x was not searched by the guard .$\rightarrow$ $\exists y$(x was accompanied by y . y was a member of the firm)].

Care must be taken, as here, to insert dots or parentheses to indicate intended grouping.

Finally, writing 'Fx' for 'x is a person that entered the building', 'Gx' for 'x was searched by the guard', 'Hxy' for 'x was accompanied by y', and 'Jy' for 'y was a member of the firm', we have:

$$\forall x[Fx \mathbin{.} -Gx \mathbin{.}\!\!\rightarrow \exists y(Hxy \mathbin{.} Jy)]$$

as the logical form of the first premise.

Instead of carrying the word 'person' explicitly through the above analysis, we might, as an alternative procedure, have limited the universe to persons. But in the present example this would have made no difference to the final symbolic form, since 'x is a person that entered the building' has all been fused as 'Fx'.

The reason for representing so long a clause as this simply as 'Fx', without further analysis, is that we know that no further analysis of it will be needed for the proposed deduction. We are assured of this by the fact that 'entered' never occurs in premises or conclusion except as applied to persons entering the building. Similarly we were able to leave 'x was searched by the guard' unanalyzed, because 'searched' never occurs except with 'by the guard'. On the other hand it behooved us to break up 'x was accompanied by some members of the firm', since accompaniment and membership in the firm are appealed to also outside this combination in the course of the premises and conclusion. In general, when we paraphrase words into logical notation and then introduce schematic letters as above, it is sound policy to *expose no more structure than promises to be needed for the proposed deduction*. This restraint not only minimizes the work of paraphrasing, but also minimizes the length and complexity of the schemata that are to be manipulated in the deduction.

Turning to the second premise, and writing 'Kx' for 'x was one of Fiorecchio's men', we get this as the obvious outward structure:

$$\exists x(Kx \mathbin{.} Fx \mathbin{.} x \text{ was unaccompanied by anyone else}).$$

It remains to paraphrase the component clause 'x was unaccompanied by anyone else'. Clearly the intended meaning is:

Anyone accompanying x was one of Fiorecchio's men,

which becomes '$\forall y(Hxy \rightarrow Ky)$', so that the second premise as a whole becomes:

$$\exists x[Kx \,.\, Fx \,.\, \forall y(Hxy \rightarrow Ky)].$$

The third premise and conclusion immediately become:

$$\forall x(Kx \rightarrow -Gx), \qquad \exists x(Kx \,.\, Jx).$$

Taking then as our four premises these three premises and the negation of the conclusion, all converted to prenex form, we proceed with the main method.

PREMISES: $\forall x \exists y(Fx \,.\, -Gx \mathrel{.\!\!\rightarrow}. Hxy \,.\, Jy)$
$\exists x \forall y(Kx \,.\, Fx \,.\, Hxy \rightarrow Ky)$
$\forall x(Kx \rightarrow -Gx)$
$\forall x \, -(Kx \,.\, Jx)$

INSTANCES: $\forall y(Kz \,.\, Fz \,.\, Hzy \rightarrow Ky)$
$\exists y(Fz \,.\, -Gz \mathrel{.\!\!\rightarrow}. Hzy \,.\, Jy)$
$Fz \,.\, -Gz \mathrel{.\!\!\rightarrow}. Hzw \,.\, Jw$
$Kz \,.\, Fz \,.\, Hzw \rightarrow Kw$
$Kz \rightarrow -Gz$
$-(Kw \,.\, Jw)$

The reader can verify that the four unquantified instances are inconsistent.

When we undertake to inject logical rigor into inferences encountered in informal discourse, we are likely to confront a second problem of interpretation over and above that of paraphrasing verbal idioms into logical notation. This second problem is that of supplying suppressed premises; and it is occasioned by the popular practice of arguing in *enthymemes*. An enthymeme is a logical inference in which one or more of the premises are omitted from mention on the ground that their truth is common knowledge and goes without saying; thus we argue:

Some Greeks are wise; for, some Greeks are philosophers,

omitting mention of the additional premise 'All philosophers are wise' on the ground that this would naturally be understood by all concerned.[2]

In everyday discourse most logical inference is enthymematic. We are constantly sparing ourselves the reiteration of known facts, trusting the

[2] Traditionally 'enthymeme' meant, more specifically, a syllogism with suppressed premise—like the above example; but it is natural, now that logic has so far outstripped the syllogism, to refer to a logical inference of any form as an enthymeme when some premises are left tacit.

listener to supply them where needed for the logical completion of an argument. But when we want to analyze and appraise a logical inference which someone has propounded, we have to take such suppressed premises into account. At this point two problems demand solution simultaneously: the problem of filling in the details of a logical deduction leading from premises to desired conclusion, and the problem of eking out the premises so that such a deduction can be constructed. Solution of either problem presupposes solution of the other; we cannot set up the deduction without adequate premises, and we cannot know what added premises will be needed until we know how the deduction is to run.

Sometimes, as in the syllogistic example above, the form of logical inference intended by the speaker suggests itself to us immediately because of its naturalness and simplicity. In such a case there is no difficulty in identifying the tacit premise which the speaker had in mind. Sometimes, on the other hand, the form of inference itself may not be quite evident, but the relevant tacit premises are already somehow in the air because of recently shared experiences. Such a case differs in no practical way from the case where all premises are explicit.

Sometimes, finally, neither the intended form of inference nor the intended tacit premises are initially evident; and in this case the best we can do is try to solve both problems concurrently. Thus we may start a tentative deduction on the basis of the explicit premises, and then, on coming to an impasse, we may invent a plausible tacit premise which would advance us toward the desired conclusion. Alternating thus between steps of deduction and supplementation of premises, we may, with luck, achieve our goal. Of course the tacit premises thus invoked must always be statements which can be presumed to be believed true by all parties at the outset; for it is only under such circumstances that a deduction using those tacit premises would give reason for belief in the conclusion. If we were to invoke as a tacit premise some statement which was (from the point of view of concerned parties) as much in need of proof as the conclusion itself, we should be guilty of what is known as *circular reasoning,* or *begging the question,* or *petitio principii;* and added conviction that might accrue to the conclusion through such argument would be deceptive. Deciding whether a statement is believed true by all parties at the outset is a task of applied psychology, but in most cases it offers no difficulty, there commonly being a wide gulf between the moot issues of an actual argument and the common fund of platitudes.

As an example of the kind of problem discussed in the foregoing paragraph, consider the explicit

PREMISES: All natives of Ajo have a cephalic index in excess of 96,
All women who have a cephalic index in excess of 96 have Pima blood

and the

CONCLUSION: Anyone whose mother is a native of Ajo has Pima blood.

Let us put these statements into logical notation, but for the present let us use obvious contractions instead of schematic letters 'F', 'G', etc., for we must keep the meanings of the words in mind in order to be able to think of relevant platitudes for use as tacit premises. The following, then, are the results of translation, supposing the universe to be comprised this time of persons:

PREMISES: $\forall x(x \text{ is nat} \rightarrow x \text{ has } 96)$,
$\forall x(x \text{ is wom} \cdot x \text{ has } 96 .\!\rightarrow x \text{ has P bl})$.
CONCLUSION: $\forall x \forall y(x \text{ is mo } y \cdot x \text{ is nat} .\!\rightarrow y \text{ has P bl})$.

When we negate the conclusion and convert it to prenex form, we get:

$$\exists x \ \exists y - (x \text{ is mo } y \cdot x \text{ is nat} .\!\rightarrow y \text{ has P bl}).$$

But we can think about it better if we subject it immediately to the obvious truth-functional simplification, thus:

$$\exists x \exists y[x \text{ is mo } y \cdot x \text{ is nat} \cdot - (y \text{ has P bl})].$$

Then we start trying for an inconsistency proof, making the most natural instantiations.

PREMISES: $\forall x(x \text{ is nat} \rightarrow x \text{ has } 96)$
$\forall x(x \text{ is wom} \cdot x \text{ has } 96 .\!\rightarrow x \text{ has P bl})$
$\exists x \exists y[x \text{ is mo } y \cdot x \text{ is nat} \cdot - (y \text{ has P bl})]$
INSTANCES: $\exists y[z \text{ is mo } y \cdot z \text{ is nat} \cdot - (y \text{ has P bl})]$
$z \text{ is mo } w \cdot z \text{ is nat} \cdot - (w \text{ has P bl})$
$z \text{ is wom} \cdot z \text{ has } 96 .\!\rightarrow z \text{ has P bl}$
$z \text{ is nat} \rightarrow z \text{ has } 96$

For further stimulation of insights we make truth-functional simplifications of the accumulating unquantified instances. The three boil down to this information:

$$z \text{ is mo } w, \qquad z \text{ is nat}, \qquad z \text{ has 96},$$
$$z \text{ is wom} \rightarrow z \text{ has P bl}, \qquad -(w \text{ has P bl}).$$

What further information is there, of a platitudinous kind, that would build this out into an inconsistency? Being mother of w, z is a woman and so, by the above conditional, has Pima blood; but then so does w, her child. So the final negation above is contradicted.

Stated as generalities, the two saving platitudes are:

$$\forall x \forall y(x \text{ is mo } y \rightarrow x \text{ is wom}),$$
$$\forall x \forall y(x \text{ is mo } y \mathbin{.} x \text{ has P bl} \mathbin{.\!\!\rightarrow} y \text{ has P bl}).$$

Putting these with the original three premises to make five, and then deriving instances from the five in our by now familiar formal fashion until a truth-functionally inconsistent lot of unquantified instances is accumulated, will be for the reader a simple bit of bookkeeping.

EXERCISES

1. Derive the inconsistent lot of instances as just now indicated.
2. Paraphrasing inward step by step, put:

   Everyone who buys a ticket receives a prize

   into symbols using '*Fxy*' for 'x buys y', '*Gy*' for 'y is a ticket', '*Hz*' for 'z is a prize', and '*Jxz*' for 'x receives z'. Then show that this implies:

   If there are no prizes then nobody buys a ticket.
3. Paraphrasing inward step by step, put:

   Every applicant who is admitted is examined beforehand

   into symbols using '*Fx*' for 'x is an applicant', '*Gxy*' for 'x is admitted at time y', '*Hxz*' for 'x is examined at time z', and '*Jzy*' for 'z is before y'. Then show that this implies:

Every applicant who is admitted is examined sometime.

4. Paraphrasing inward step by step, put:

   There is a painting which is admired by every critic who admires any paintings at all

   into symbols using '*F*' for 'painting', '*G*' for 'critic', and '*H*' for 'admires'. Then add the further premise:

   Every critic admires some painting or other

   and show that these premises imply:

   There is a painting which all critics admire.

5. Assume that I like anyone who laughs at himself but detest anyone who laughs at all his friends. Show that these premises, and a platitudinous further one, together imply that if anyone laughs at all his friends then someone is no friend of himself. (Quimby)

[3]

32 COMPLETENESS

In proving the completeness of the main method we shall need what I call the *law of infinite conjunction*. It is substantially what has been known in topology as König's infinity lemma or Brouwer's fan theorem, and elsewhere as the compactness theorem. In the form suited to our purposes, it says that *an infinite class of truth-functional schemata is consistent if each of its finite subclasses is*.

Figuratively speaking the law says that an infinite conjunction of truth-functional schemata is consistent if each finite conjunction of those schemata is separately consistent. This formulation gives the law its name, but is only figurative, since conjunctions are expressions, and expressions, literally speaking, are only finite strings of marks. The literal

[3] All the rest of Part III is bracketed for omission for purposes of a short course.

statement of the law speaks rather of an infinite class, as above. A class of truth-functional schemata is called consistent if there is an assignment of truth values to sentence letters that makes all the schemata in the class come out true.

To see why the law of infinite conjunction is true, assume some infinite class K of truth-functional schemata and assume that each K-conjunction (each finite conjunction of members of K) is consistent. Let us represent the sentence letters in these schemata as 'p_1', 'p_2', etc.

Definition. A given assignment of truth values to one or more sentence letters will be called *innocuous* (so far as K is concerned) if it conflicts with no K-conjunction, causing it to resolve to $\perp$ or to an inconsistency.

Definition. t_1 is $\top$ or $\perp$ according as assignment of $\top$ to 'p_1' is innocuous or not.

Lemma 1. Assignment of t_1 to 'p_1' is innocuous.

Proof. If assignment of $\top$ to 'p_1' conflicted with some K-conjunction and assignment of $\perp$ to 'p_1' conflicted with one too, then the conjunction of the two conjunctions would be an inconsistent K-conjunction outright, contrary to our hypothesis on K. So one or other of $\top$ and $\perp$ must be an innocuous assignment to 'p_1'. Either way, t_1 is then innocuous by its definition.

Definition. t_2 is $\top$ or $\perp$ according as assignment of t_1 to 'p_1' and $\top$ to 'p_2' is innocuous or not.

Lemma 2. Assignment of t_1 to 'p_1' and t_2 to 'p_2' is innocuous.

Proof. If assignment of t_1 to 'p_1' and $\top$ to 'p_2' conflicted with some K-conjunction and assignment of t_1 to 'p_1' and $\perp$ to 'p_2' conflicted with one also, then mere assignment of t_1 to 'p_1' would conflict with the conjunction of the two conjunctions, contrary to Lemma 1. So either the assignment of t_1 to 'p_1' and $\top$ to 'p_2' or the assignment of t_1 to 'p_1' and $\perp$ to 'p_2' is innocuous. Either way, then, Lemma 2 holds, by virtue of the definition of t_2.

Definition. t_3 is $\top$ or $\perp$ according as assignment of t_1 to 'p_1', t_2 to 'p_2' and $\top$ to 'p_3' is innocuous or not.

Lemma 3. Assignment of t_1 to 'p_1', t_2 to 'p_2', and t_3 to 'p_3' is innocuous.

Proof from Lemma 2 like that of Lemma 2 from Lemma 1.

We have in the definitions the beginning of a series whose continuation is evident. Truth values t_1, t_2, t_3, t_4, . . . are thereby defined which are designed for assignment to 'p_1', 'p_2', 'p_3', 'p_4', ... Continuing the lemmas to Lemma i, for any desired i, we show that assignment of t_1, t_2,

. . . , t_i to 'p_1', 'p_2', . . . , 'p_i' is innocuous. Consider, then, any schema S in K, and take i large enough so that 'p_1', 'p_2', . . . , 'p_i' exhaust the sentence letters in S. Then assignment of t_1, t_2, . . . , t_i to those letters resolves S to $\top$ or $\bot$; but, by Lemma i, not to $\bot$; so to $\top$. But S was any member of K. So assignment of t_1, t_2, . . . to 'p_1', 'p_2', . . . makes all members of K come out true. So K is consistent, Q.E.D.

I pause for a brief digression on constructiveness. Once the infinite class K has been specified, in whatever fashion, the truth values t_1, t_2, etc. are uniquely fixed by the definitions above. But this does not mean that, even knowing K, we can find out whether t_1 is $\top$ or $\bot$, nor whether t_2 is $\top$ or $\bot$, and so on. Even if we have a decision procedure for membership in K, we may have none for whether the assignment of $\top$ to 'p_1' would conflict with some or no conjunctions of members of K. On this account the proof of the law of infinite conjunction is not what is called a *constructive* proof. It proceeded by showing the existence of certain suitable objects (in this case t_1, t_2, etc.) without saying just which objects they are (e.g., whether t_1 is $\top$ or $\bot$). Much that is proved in mathematics is insusceptible to constructive proof; much else is proved constructively.

We turn now to the theorem of completeness of the main method. It says of any combination of (one or more) quantificational schemata that *either* they come out jointly true under some interpretation in a nonempty universe *or* some truth-functionally inconsistent combination of unquantified schemata can be derived from them by the familiar operations: convert to prenex form, reletter bound variables that match free ones, and then apply UI and EI, using only new instantial variables in EI.

Suppose then some premises, converted to prenex form and relettered as required. A rigid routine will be described for progressively instantiating them. It will be such as to assure that each existential line (each line beginning with an existential quantifier) gets instantiated once, and that each universal line gets instantiated with each free variable that ever turns up.

Begin with a great wave of EI: instantiate each existential line once, using a variable not yet used at all. If new existential lines emerge in the process, instantiate them too. Then follow with a great wave of UI: instantiate each universal line, whether old or emergent, with each variable that is already free in the proof. (Some may be free in premises; others will have issued from EI.) Then another wave of EI, instantiating any further existential lines brought by the wave of UI. Then a double wave of UI: the old universal lines get instantiated with the newly added free variables, and any new universal lines get instantiated with all the

variables now free in the proof. Then another wave of EI; then another double wave of UI; and so on.

In the special case where all premises are closed and universal, this pump-priming operation is necessary: perform UI using a new variable.

Here is an example of the rigid routine. I shall prolong the alphabet of variables with help of accents: 'z'', 'z''', etc.

PREMISES:	$\forall w \exists x(Fz \vee Gwx)$	
	$\forall w \exists x \forall y(Gwx \cdot -Fy)$	
INSTANCES:	$\exists x(Fz \vee Gzx)$	(2d wave; 1st was empty)
	$\exists x \forall y(Gzx \cdot -Fy)$	(2d wave)
	$Fz \vee Gzz'$	(3d wave)
	$\forall y(Gzz'' \cdot -Fy)$	(3d wave)
	$\exists x(Fz \vee Gz'x)$	(4th wave, 1st part)
	$\exists x(Fz \vee Gz''x)$	(4th wave, 1st part)
	$\exists x \forall y(Gz'x \cdot -Fy)$	(4th wave, 1st part)
	$\exists x \forall y(Gz''x \cdot -Fy)$	(4th wave, 1st part)
	$Gzz'' \cdot -Fz$	(4th wave, 2d part)
	$Gzz'' \cdot -Fz'$	(4th wave, 2d part)
	$Gzz'' \cdot -Fz''$	(4th wave, 2d part)
	$Fz \vee Gz'z'''$	(5th wave)
	$Fz \vee Gz''z''''$	(5th wave)

and so on. The reader should examine this for conformity to the rigid routine.

Some examples terminate, with or without inconsistency, and some do not. In the unending cases, new free variables get introduced by EI without end; we may suppose them generated systematically by accentuation as above.

For each universal line, in any event, and each free variable that will ever turn up, we may be sure that that line will eventually be instantiated with that variable. For, each wave of EI, importing new variables, is followed by a wave of UI that uses them all. Moreover, each existential line is bound eventually to be instantiated once. For, under the above routine an existential line will be instantiated as soon as all unfinished business has been cleared up; and that business is limited.

Let $A_1, A_2, \ldots$ be, in an arbitrary order, all the schemata obtainable by applying term letters of the premises to variables that are free in the perhaps infinite suite of instances generated by the above routine. Thus in

the above example $A_1, A_2, \ldots$ might be 'Fz', 'Gzz', 'Fz'', 'Gzz'', '$Gz'z$', '$Gz'z'$', 'Fz''', and so on. Whether there is an end to $A_1, A_2, \ldots$ will depend on whether there is an end to the instances; in some examples there is an end and in some not. In any event the unquantified instances will be truth functions of A_1, A_2, etc.

What was to be shown was that either the premises come out true under some interpretation in a nonempty universe or some finite set of unquantified instances is truth-functionally inconsistent. This will be shown to hold even when we consider only the instances generated under the rigid routine above.

I shall prove the theorem in this form: if each finite set of unquantified instances is truth-functionally consistent, then the premises are true under an interpretation in a nonempty universe—and a universe, in fact, of positive integers.

So suppose each finite set of unquantified instances is consistent. Then by the law of infinite conjunction (using A_1, A_2, etc. in place of 'p_1', 'p_2', etc. in that law) there is an assignment $\mathfrak{A}$ of truth values $t_1, t_2, \ldots$ to $A_1, A_2, \ldots$ that makes all the unquantified instances true. Now let the universe consist of as many of the positive integers as there are free variables in the instances—all the positive integers if the variables are unending. Interpret the variables, say in order of first appearance, as naming 1, 2, etc. Then interpret each term letter as true of just the integers, or pairs, etc., that the assignment $\mathfrak{A}$ declares it true of.

Thus consider again our example. An assignment $\mathfrak{A}$ of truth values to 'Fz', 'Gzz', 'Fz'', 'Gzz'', etc., that makes all the infinitely many unquantified instances in that example come out true, is this: $\bot$ to all of 'Fz', 'Fz'', etc., and $\top$ to all of 'Gzz', 'Gzz'', '$Gz'z$', '$Gz'z'$', 'Gzz''', '$Gz'z''$', '$Gz''z$', etc. So, taking z as 1, z' as 2, z'' as 3, etc., we proceed to interpret 'F' as true of no integers and 'G' as true of all pairs of integers. Thus this example turns out unusually trivial.

The plan is, in short, so to interpret the term letters as to grant A_1, A_2, etc. the respective truth values t_1, t_2, etc. In thus realizing the assignment $\mathfrak{A}$, the interpretation makes all the unquantified instances come out true.

But then any singly quantified instance comes out true too, if its quantifier is existential. For, we insured that each such line (indeed each existential line) gets instantiated; its instance will imply it, since '$\ldots z \ldots$' implies '$\exists x(\ldots x \ldots)$'; and its instance will be true, being unquantified.

Also any singly quantified instance comes out true if its quantifier is universal. For, call it '$\forall x(\ldots x \ldots)$'. We insured that any such line gets instantiated with each free variable that ever turns up. But all these

U-instances '. . . z . . .', '. . . z' . . .', etc. are true, being unquantified, and moreover the integers 1, 2, etc. that are the interpretations of these variables exhaust the universe; so '. . . x . . .' comes out true of everything in the universe, and thus '$\forall x(\ldots x \ldots)$' comes out true.

From the truth of all unquantified instances we have been able to infer the truth of all singly quantified instances. From this result it follows similarly that all doubly quantified instances come out true. Repeating this reasoning as many times as a premise has quantifiers, we show finally that each premise comes out true.

The schemata are thus found to come out true under an interpretation in a nonempty universe. The universe was not empty thanks to the pump-priming operation, which assured at least one free variable and thus one integer.

HISTORICAL NOTE: The discovery and proof of the completeness of a proof procedure for the logic of quantification dates from work of Skolem, Herbrand, and Gödel, 1928–1930. It became explicit in Gödel. The proof procedures there concerned differ from ours, but this difference is of little moment; a completeness proof for one method of quantification theory can be adapted fairly easily to others. In the adaptation above I have depended partly on Gödel's original argument and partly on a variant due to Dreben.

EXERCISES

1. Continue the illustrative instantiations above, adding a dozen more.

2. Where the members of K are '$p_1 \rightarrow \bar{p}_2$', '$p_2 \rightarrow \bar{p}_3$', '$p_3 \rightarrow \bar{p}_4$', and so on without end, determine t_1, t_2, t_3, etc.

3. The definition of t_i that was given, for each i, began: 't_i is $\top$ or $\bot$ according as . . .'. Try switching this to read 't_i is $\bot$ or $\top$ according as . . .'. Can Lemmas 1, 2, 3, etc. still be proved for these changed values? Check each step.

33
LÖWENHEIM'S THEOREM

The completeness proof serves at the same time to establish also an important older theorem due to Löwenheim (1915): *If a quantificational schema comes out true under an interpretation in a nonempty universe at all, it comes out true under an interpretation in the universe of positive integers.* The reasoning is as follows. The schema cannot engender an inconsistency under the main method, since that method is sound (Chapter 30). By the completeness proof above, then, the schema (if it is prenex) will come out true under some interpretation in a nonempty universe of positive integers. But then, by the argument in the middle of Chapter 18, it will come out true also under some interpretation in any containing universe—hence in the universe of all positive integers. Löwenheim's theorem thus holds for prenex schemata; and then it carries over to the others by virtue of their convertibility to prenex form.

There is no difference between affirming Löwenheim's theorem for single schemata and affirming it for finite classes of them, since finitely many schemata can be joined in conjunction. But its extension to infinite classes of schemata is a genuine extension, and was made by Skolem in 1920. It is the Löwenheim-Skolem theorem: *If all of a class of quantificational schemata come out true together under an interpretation in a nonempty universe, they come out true together under some interpretation in the universe of positive integers.*

The argument remains much the same as in the completeness proof. We suppose given, now, an infinite consistent class C of prenex schemata. Let C_1 be C plus all direct E-instances of its schemata, plus E-instances of those instances, and so on, with a new instantial variable for each. Let C_2 be C_1 plus instances of all its universal schemata, the instantial variables being all the free variables of C_1. Let C_3 be C_2 plus E-instances; and so on. It is just the rigid routine again, except that this ceases now to be a routine of performances and becomes rather a definition of an ascending series of infinite classes.

Finally consider all the unquantified instances that are reached in any of these classes $C_1, C_2, \ldots$. If a finite lot of these were truth-functionally

inconsistent, then the schemata in C that were their ultimate sources could be proved jointly inconsistent by the main method. But the whole original class C was consistent, by hypothesis. So we conclude that each finite lot of unquantified instances is truth-functionally consistent. So then are they all, by the law of infinite conjunction. From these unquantified instances, then, get $A_1, A_2, \ldots$ as in Chapter 32, and $\mathfrak{A}$, and the interpretation of the term letters in a universe of positive integers.

An importance of the Löwenheim-Skolem theorem, in its coverage of infinite classes of schemata, may be seen as follows. Consider any nonempty universe U and any assortment of terms, all interpreted in that universe. Consider, further, the whole infinite totality of truths, known and unknown, that are expressible with help of those terms together with the truth functions and quantification over U. Then the Löwenheim-Skolem theorem assures us that there is a reinterpretation of the terms, in the universe of positive integers, that preserves the whole body of truths.

E.g., taking U as the universe of real numbers, we are told that the truths about real numbers can by a reinterpretation be carried over into truths about positive integers. This consequence has been viewed as paradoxical, in the light of Cantor's proof that the real numbers cannot be exhaustively correlated with integers. But the air of paradox may be dispelled by this reflection: whatever disparities between real numbers and integers may be guaranteed in those original truths about real numbers, the guarantees are themselves revised in the reinterpretation.

In a word and in general, the force of the Löwenheim-Skolem theorem is that the narrowly logical structure of a theory—the structure reflected in quantification and truth functions, in abstraction from any special terms—is insufficient to distinguish its objects from the positive integers.

To say that a schema or class of schemata comes out true under some interpretation in the universe of positive integers is not to say that we can write out suitable interpretations of 'Fxy', 'Gx', etc. as formulas in arithmetical notation. Our proof of Löwenheim's theorem assured us that for any consistent schema there *is* a true interpretation in the universe of positive integers, but it gave us no way of finding and phrasing it. Our proof used the completeness theorem, the proof of which appealed, after all, to the nonconstructive t_1, t_2, etc. Nor does our proof assure us that the interpretation, even if it were found, could be phrased in a purely arithmetical vocabulary—the vocabualry of 'plus', 'times', 'power', 'equals', and the like.

There is, however, a constructive improvement on Löwenheim's theorem due to Hilbert and Bernays[4] that tells us all this: how to find and

actually write out, for any consistent schema, a true interpretation in the universe of positive integers—and in purely arithmetical notation at that. In fact a vocabulary amounting to 'plus', 'times', 'equals', and the quantifiers and truth functions is sufficient—the vocabulary, as one says, of *elementary number theory*.

This result appears the more remarkable when we read, as we shall in the next chapter, that there can never be a fully general technique for establishing consistency. How then can Hilbert and Bernays give us a general rule for writing out a true interpretation of a consistent schema? What they give is a general rule for writing out, for *any* schema, an arithmetical interpretation which can be depended upon to come out true in case the schema does happen to be consistent. This strengthening cannot be managed for the Löwenheim-Skolem theorem, with its coverage of infinite classes of schemata, but it works for Löwenheim's.

The proof of this Löwenheim-Hilbert-Bernays theorem is beyond the scope of this book. It will be worthwhile, however, to notice a certain bearing that it has on the very notion of interpretation.

In Chapters 18 and 27 there was no outright statement of what an interpretation of a term letter *is*, but only a statement of what it means to have given one: it means having settled what members of the universe, or what pairs, etc., the letter is to be counted as true of. One way of settling this is by supplying an abstract (Chapter 26), but there are other more discursive ways. Also there are degrees of settling, and, therewith, degrees of having interpreted; we interpret the term letter *insofar* as we settle what to take it as true of.

To face it, then: what is an interpretation? Is it an expression, namely the abstract that we are prepared to substitute for the term letter? Or is it a class, namely the class of the things that the term letter is taken to be true of? *Prima facie*, the choice between these two alternatives should make a material difference; for there is no assurance that each class of objects in our universe is specifiable as the extension of some term or open sentence in our language. It depends on our choice of universe and also on how rich a vocabulary we assume there to be at our disposal over and above our logical notations of truth functions and quantification. If the classes thus outrun the available expressions, then we may expect there to be a difference, e.g., in the concept of validity, according as we define it as truth under all interpretations in the class sense or as truth under all

[4] Volume 2, pp. 234–253. For a strengthened version, more simply proved, see Kleene, 1952, p. 394, Theorem 35.

interpretations in the term sense. Moreover, it has been an accepted tenet of classical set theory from Cantor onward that the classes do outrun the expressions.

But now a remarkable thing about the Löwenheim-Hilbert-Bernays theorem is that it cuts through all this. However extravagantly the classes may outrun the expressions, this theorem assures us that when we define validity and consistency it is indifferent whether we talk of all and some interpretations in the sense of classes or in the sense of terms. The theorem assures us that all those extra classes will be indifferent to validity and consistency, in this sense: if a schema is fulfilled (or falsified) by some unspecifiable interpretation involving nameless classes, it is *also* fulfilled (or falsified) by some other interpretation that can be written in the notation of arithmetic.

The Löwenheim-Hilbert-Bernays theorem enables us, for that matter, simply to bypass the term 'interpretation' and talk directly of substitution. A schema is valid if all statements obtainable from it by substitution are true, and it is consistent if some of them are true. The theorem assures us that we can retreat to this homely formulation with impunity as long as the vocabulary which we suppose available for such substitutions is not so grotesquely impoverished as to be inadequate to elementary number theory.

There is philosophical comfort in the assurance that we can talk of logical validity and consistency without appealing to a limitless realm of abstract objects called classes. We feel that in talking of substitution of expressions we still keep our feet on the ground.

A related point could be made on the basis directly of the completeness theorem, even independently of Löwenheim, Hilbert, and Bernays. That theorem already furnishes formulations of validity and consistency that say nothing of classes—nor even of substitution of predicates, for that matter. A schema is valid if, from the prenex form of its negation, a truth-functional inconsistency is derivable by the main method.

34
DECISIONS AND THE UNDECIDABLE

It is indifferent to the main method whether we write our premises with or without initial existential quantifiers. This is not to be wondered at, for we noted early in Chapter 25 that initial existential quantifiers were indifferent to consistency. This indifference is reflected in the rigid routine of Chapter 32, since that routine provides that an existential quantifier, if not preceded by a universal, gets instantiated once for all; we never get back to it. The E-instance with its instantial variable could as well have been the premise to begin with.

It is the existential quantifiers preceded by universals that make proofs run long. Such an existential quantifier returns to the surface repeatedly as those universal quantifiers get repeatedly and variously instantiated. Each time it surfaces it gets instantiated again, since it initiates a line each time that is altered by having received different instantial variables through the new UI steps. Each of these EI steps, then, furnishes a new variable that is grist for further UI, and in this way we may go on grinding out instances without end. The unending example at the middle of Chapter 32 should be looked at with this in mind.

When all premises are such that no existential quantifiers are preceded by universals, our rigid routine of Chapter 32 always grinds gratifyingly to a halt. There is just the fixed initial supply of what are in fact or in effect free variables, and these are all that get used in UI. The total number of instances allowed by the rigid routine, for each such premise, will be just the number of universal quantifiers in that premise times the number of free (or existential) variables in the premises altogether. Indeed, as already remarked, the existential quantifiers, being initial here, could have been dropped in the first place; our premises are nothing more than *pure universals*. We can check the totality of their unquantified instances for truth-functional consistency, confident of an affirmative or negative answer. We are not then in the position of just checking more and more unquantified instances for inconsistency with never a notion of when to give up; we have a decision procedure.

In Chapter 29 we saw a decision procedure for validity of pure existentials. This already, implicitly, was a decision procedure for consistency of pure universals. For a schema is inconsistent if and only if its negation is valid; and the negation of a pure universal becomes a pure existential under the rules of passage (9) and (10) of Chapter 23. What is interesting in the decision procedure for consistency that has just now emerged from the main method is only the particular form in which it has appeared, and its relation to the main method. It does suggest that in converting schemata to prenex form we should favor existential quantifiers if we expect to apply the main method, just as we did the opposite when trying for pure existentials. Actually we can effortlessly switch course when we see fit, thanks to the aforementioned rules of passage (9) and (10) of Chapter 23.

I have alluded twice to the impossibility of a decision procedure for quantificational schemata generally. The proof, due to Church and Turing in 1936, exceeds the scope of this book,[5] but some remarks are called for regarding its significance. It does not preclude mechanical proof procedures; the rigid routine of Chapter 32 is one. Any proof procedure can in principle be got down to a mechanical routine in at least this silly way: just scan all the single typographical characters admissible in the proof procedure in question, then all possible pairs of them, then all strings of three, and so on until you get to the proof. The usual dependence on luck and strategy in finding proofs is merely a price paid for speed: for anticipating, by minutes or weeks, what a mechanical routine would eventually yield.

Where a complete proof procedure differs from a decision procedure, then, is not in being less mechanical. The difference is just that the proof procedure is not a yes-or-no affair; it does not deliver negative answers. Failure to reach an inconsistency proof after a large number of steps, however systematically and mechanically programmed, does not show consistency. A proof procedure is only half of a decision procedure.

When there is both a proof procedure and a disproof procedure (for validity or consistency or any other property), there is also a decision procedure.[6] The two halves can be joined as follows. First we would mechanize both the proof procedure and the disproof procedure, in the silly way if not otherwise. Then, to decide whether a formula had the property in question (validity or consistency or whatever), we would set

[5] For a fairly readable rendering see my *Selected Logic Papers*, pp. 212–219.

[6] This observation is due to Kleene, 1943. Or see Kleene, 1952, p. 284, Theorem VIc.

one man or machine in quest of a proof and another in quest of disproof and await the eventual answer from one of them.

In particular, then, knowing as we do from Chapter 32 that there is a complete procedure for disproving the consistency of a quantificational schema, and from Church and Turing that there is no decision procedure, we conclude that *there is no complete procedure for proving quantificational schemata consistent*. Commonly we prove a schema consistent by producing a true interpretation; and this is quite the right way, as long as we can find truths to the purpose and show them true. So what we have to recognize is that no sound proof procedure can be so strong as to afford proofs of statements illustrative of all consistent quantificational schemata.

Meanwhile there is an obvious procedure for deciding whether a schema comes out true under all or some or no interpretations in a universe of given finite size n. Just transcribe the universal and existential quantifications as n-fold conjunctions and alternations, as noted in Chapter 27, and then check for truth-functional validity and consistency. It follows that there is a complete procedure (no longer a decision procedure, but still a proof procedure) for showing that a schema is *finitely consistent*—i.e., that it comes out true under at least one interpretation in at least one nonempty finite universe. This proof procedure consists simply in alleging an adequate size n of the universe, and leaving the reader to check it by the truth-functional method.

Thus an important contrast emerges between finite and infinite universes: there is a complete proof procedure for finite consistency, but none for consistency. The schemata that are consistent but not finitely so are the stubborn kind. They may be called *infinity schemata*. An easily recognizable example is:

$$(1) \qquad \forall x[-Fxx \,.\, \forall y \forall z(Fxy \,.\, Fyz \,.\!\rightarrow Fxz) \,.\, \exists w\, Fxw].$$

It requires an infinite universe. For, take anything x_1. By the last part of (1) we have Fx_1x_2 for some x_2; also Fx_2x_3 for some x_3; and so on without end. Then, by the transitivity affirmed in the middle of (1), we get also Fx_1x_3, Fx_1x_4, Fx_2x_4, Fx_1x_5, and so on. Moreover, by the '$-Fxx$' in (1), these things x_1, x_2, x_3, . . . are all distinct. Thus (1) is not finitely consistent. Yet it is consistent; it comes out true in the universe of integers when 'F' is interpreted as '$<$'.

Here is another infinity schema, briefer but less readily recognized:

$$\forall x \exists y \forall z (Fxy \,.\, -Fxx \,.\, Fyz \rightarrow Fxz).$$

From a 1933 paper by Gödel it is known that this prefix '$\forall x \ \exists y \ \forall z$' is the simplest that a prenex infinity schema can have.

There is no complete proof procedure for showing schemata to be infinity schemata. For, if there were, we could add it to our complete proof procedure for finite consistency, and get a complete proof procedure for consistency.

The infinity schemata constitute a stubborn class which is the common part AB of two stubborn classes: A, the class of all consistent schemata, and B, the class of all finitely inconsistent schemata (schemata true under no interpretations in nonempty finite universes). We have observed that there is no complete proof procedure for membership either in A or in AB. The same can be said of B; this follows from a theorem of Trachtenbrot, according to which there is no complete proof procedure for finite validity.

Taken in combination with the Löwenheim-Hilbert-Bernays theorem (Chapter 33), the impossibility of a complete proof procedure for A has a startling further consequence: Gödel's theorem of the impossibility of a complete proof procedure for elementary number theory. If we could prove each true sentence of elementary number theory, then we could prove the consistency of each consistent quantificational schema by proving a truth of elementary number theory illustrative of that schema.

This is not the historical route to Gödel's theorem. I have appealed here to the impossibility of a complete proof procedure for consistency. This I inferred from the Church-Turing theorem that there is no decision procedure for quantification validity. But the proof of the Church-Turing theorem, omitted here, uses the essential trick of Gödel's proof of his own theorem; and Gödel's came first, in 1931.

Gödel's way was to argue directly to the incompletability of elementary number theory, by showing how, for any given proof procedure P for elementary number theory, a statement S_P of elementary number theory can be constructed which will be true if and only if it is not provable by the procedure P. Either S_P is provable, in which case it is false and so the general proof procedure P is discredited, or else S_P is true and not provable, in which case the proof procedure P is incomplete.

In broadest outlines, Gödel's construction of S_P is as follows. It is easy to assign integers in a systematic way to all the finite strings, however long, of signs of a given alphabet. If the alphabet has just nine signs, we can assign them the integers from 1 to 9 and then we can get the integer

for any string of signs by concatenating the corresponding digits into a long numeral. If the alphabet has more than nine signs, we can easily adapt this method or choose another. Suppose this done for the notation of elementary number theory. Thus each sentence in that notation has its so-called *Gödel number*. Gödel then shows that, given P, it is possible within the notation of elementary number theory to formulate an open sentence, '. . . x . . .' say, which is true of any number x if and only if x is the Gödel number of a statement provable by P. If we put for 'x' in '$-(\ldots x \ldots)$' the actual numeral designating some particular number n, clearly the resulting statement, schematically '$-(\ldots n \ldots)$', will be true if and only if that chosen n is not the Gödel number of a statement provable by P. But Gödel shows that n can be so chosen as to turn out to be the Gödel number of a statement equivalent to '$-(\ldots n \ldots)$' itself. The statement produced by that sly choice of n is the sought S_P, true if and only if not provable by P.[7]

When we are concerned with complete proof procedures and decision procedures for the truth of statements, as in elementary number theory, rather than for properties of schemata, the difference between the two disappears—at any rate if our vocabulary includes negation. For, a proof procedure for the truth of statements carries with it a disproof procedure: you disprove a statement by proving its negation. So, by the zigzag argument lately attributed to Kleene, a complete proof procedure for the truth of statements assures a decision procedure. Still a complete proof procedure for validity or consistency or inconsistency of schemata remains far short of a decision procedure, simply because a schema can be consistent without its negation's being inconsistent, and a schema can fail of validity without its negation's being valid.

Thus, though the impossibility of a complete proof procedure for elementary number theory comes as a shock, the impossibility of a decision procedure for elementary number theory says as much. This is a curious point, since the impossibility of a decision procedure seems less surprising than the other. Various unsolved problems of long standing, after all—notably the celebrated one of Fermat—can be formulated within the notation of elementary number theory, and a decision procedure for that domain would have made a clean sweep of them all.

Presburger and Skolem have shown that when elementary number theory is further limited to the extent of dropping multiplication and

[7] Fuller details of Gödel's argument may be found not only in his 1931 paper but also in his 1934 English presentation, and in Carnap's *Logical Syntax* (esp. pp. 129–134), and in Chapter 7 of my *Mathematical Logic*.

keeping just addition, or vice versa, the resulting theory does admit of a decision procedure. What is more surprising, Tarski has shown that the elementary algebra of real numbers likewise admits of a decision procedure. The notation of this elementary algebra is precisely the same as that described above for elementary number theory, including both addition and multiplication; the only difference is that the variables are construed now as referring to real numbers generally rather than just to whole ones. Despite the seemingly greater complexity of its subject matter, elementary algebra is completable and mechanically decidable while elementary number theory is not.

Now if in view of Gödel's result our knowledge about number is subject to unexpected limitations, the very opposite is true of our knowledge about such knowledge. One of the few things more surprising than the incompletability of elementary number theory is the fact that such incompletability can actually have become known to us. Gödel's result brings a new branch of mathematical theory to maturity, a branch known as "metamathematics" or proof theory, whose subject matter is mathematical theory itself.

Much remains to be discovered regarding the limits of completable theories, and the essential structural features which set such theories apart from those which are incompletable. But the reader who would acquire the key concepts and techniques of this crucial new field of study, metamathematics or the theory of proof, must look beyond the bounds of this logic book. See Tarski, et al., *Undecidable Theories;* Hilbert and Bernays; Kleene; Smullyan; Davis; Rogers; Shoenfield.

EXERCISES

1. Can there be a complete disproof procedure for elementary number theory? Or a decision procedure adequate to all but a finite number of the statements in the notation of elementary number theory? Justify your answers.

2. From Trachtenbrot's theorem how does it follow that there can be no complete proof procedure for membership in B?

35
FUNCTIONAL NORMAL FORMS

Getting all existential quantifiers out in front has, we saw, two advantages under our main method: the minor advantage that we can just forget those quantifiers, and the major advantage that the proof procedure becomes a decision procedure. It also has a third advantage: the remaining quantifiers, being uniformly universal, can be instantiated simultaneously rather than in sequence.

Now there is a devious device that enables us to enjoy this third benefit even when existential quantifiers can not all be got to the front and a decision procedure is not to be hoped for. To see the fundamental idea, consider '$\forall x\ \exists y\ Fxy$'. It says that for any object x there is an object y, possibly different for different choices of x, such that Fxy. Roughly speaking therefore it says there is a *function*—a way of picking, for each object x, a dependent object y_x—such that $\forall x\ Fxy_x$. The gain in this rephrasing is that our new tacit existential quantifier—'there is a function'—comes before '$\forall x$' instead of after it as '$\exists y$' did. Coming thus initially, it can be left tacit in inconsistency proofs.

Such is the thought behind what I call the *functional normal form* of a prenex quantificational schema. Delete all existential quantifiers and attach, to the recurrences of the variable of each such quantifier, subscripts listing the variables of all the universal quantifiers that preceded that existential quantifier. Thus the functional normal form of '$\forall x\ \exists y\ Fxy$' is '$\forall x\ Fxy_x$'. The functional normal form of:

$$\exists x\ \forall y\ \forall z\ \exists w\ \forall u\ \exists v\ Fxyzwuv$$

is:

$$\forall y\ \forall z\ \forall u\ Fxyzw_{yz}uv_{yzu}.$$

The account of this notation in terms of the existence of functions is not an account to rest with, as in any sense a proof of equivalence of a schema to its functional normal form. The account slurred over points of set

theory that I shall not go into—the nature and existence of functions and, what is more serious, something known as the axiom of choice. None of this is needed to show the soundness or completeness of the proof procedure that I shall present in terms of functional normal forms; these virtues will be brought out rather by a direct collation of this proof procedure with the main method. The point of the talk of functions was just to give a feeling for the notation.

We turn now to illustration and closer examination of the method of proving inconsistency by functional normal forms. Justfication will come afterward.

Let us go back to the premises (1)–(3) in Chapter 30. A rule of the method of functional normal forms is that before converting premises to functional normal form we must reletter all existential quantifiers so that their variables are distinct from one another and from variables that are free in any premise. The premises (1)–(3) are already in order on this score. Putting them into functional normal form and deriving instances purely by UI, we get this new inconsistency proof:

PREMISES: $\forall y\, Fuy$

$\forall x\, \forall y\, -(Fv_xy \,.\, Fyx)$

$\forall x\, \forall y(Fxy \rightarrow\!: -Fxw_{xy} \,.\, Fw_{xy}y \;.\vee.\; Fyw_{xy} \,.\, Fw_{xy}x)$

INSTANCES: Fuv_u

$Fuv_u \rightarrow\!: -Fuw_{uv_u} \,.\, Fw_{uv_u}v_u \;.\vee.\; Fv_uw_{uv_u} \,.\, Fw_{uv_u}u$

Fuw_{uv_u}

$-(Fv_uw_{uv_u} \,.\, Fw_{uv_u}u)$

These four instances, like the four unquantified instances (7), (10), (11), and (12) in the rendering in Chapter 30, are truth-functionally inconsistent. But it is characteristic of the method of functional normal forms that the only instances that appear at all are unquantified. The instances come directly from the premises by multiple UI, and it is important to note that the universal variables are instantiated in subscript positions as well as elsewhere. The expressions instantiating them are no longer just simple variables, but any expressions belonging to what I call the *lexicon* of the premises. To get the lexicon we look to the premises in functional normal form and pick out, first, the simple free variables—in this case just '*u*'. They belong to the lexicon. Then we take also the function letters, i.e., the letters that carry subscripts; each function letter, taken together with any members of the lexicon in place of its subscripts, constitutes a member in turn of the lexicon. Thus the lexicon of the premises above

comprises 'u', 'v_u', 'w_{uu}', 'v_{v_u}', 'w_{v_uu}', 'w_{uv_u}', '$w_{v_uv_u}$', '$v_{w_{uu}}$', and so on. In proving inconsistency there is never any need to instantiate with anything but lexicon. (If there is no free variable, start lexicon with 'x'.)

This method of inconsistency proof is an extension of one that was presented by Skolem in 1928. Essentially the difference is just that his served only to prove the inconsistency of a single prenex premise rather than a combination of them.

The method is advantageous not merely in allowing us to perform several UI steps at once. A more important advantage is, in a mountaineering methaphor, that it eliminates the need of advance bases. Under the main method we would move perhaps from a premise '$\forall x\ \exists y\ \forall z\ Fxyz$' to '$\exists y\ \forall z\ Fvyz$' to '$\forall z\ Fvwz$' to '$Fvwv$', recording and saving the two intermediate lines, and then later we might recur to one of these intermediate lines, '$\forall z\ Fvwz$', to get to '$Fvww$'. Under the method of functional normal forms, on the other hand, we pass from the premise direct to '$Fvwv$' and again direct to '$Fvww$' in two independent and uninterrupted sorties, with no need of the branching point '$\forall z\ Fvwz$' as a way station. Under this method the premise has really become '$\forall x\ \forall z\ Fxy_xz$', of course, and the two end results '$Fvwv$' and '$Fvww$' have become, strictly speaking, 'Fvy_vv' and 'Fvy_vy_v'.

To relate this method to the main method, take the main method in the relaxed form in which it was last contemplated, viz., with initial existential quantifiers dropped from all premises. The only existential quantifiers in the premises are those buried under universal quantifiers. Now consider the route, according to the main method, from one of these premises—say again '$\forall x\ \exists y\ \forall z\ Fxyz$'—to each of its unquantified instances. Why can we not cover each such route in one leap? Simply because the choice of instantial variables for the EI step is not free. Whenever the '$\forall x$' is instantiated anew, to yield a novel intermediate line, the ensuing EI step demands a fresh variable. When on the other hand the routes from the premises to two unquantified instances instantiate '$\forall x$' identically but diverge ultimately on '$\forall z$', we are free to instantiate '$\exists y$' identically; under the rigid routine of Chapter 32, indeed, we would be bound to. The main method handles these dependences by establishing an advance base or way station like '$\forall z\ Fvwz$', we noted, from which to proceed now to '$Fvwv$' and now to '$Fvww$'. The alternative, instant instantiation, requires automatic bookkeeping, calculated to allow two unquantified instances of '$\forall x\ \exists y\ \forall z\ Fxyz$' to have any variables in place of 'x' and 'z', but to make them agree or disagree on the variable for 'y' according as they agree or disagree on that for 'x'. (I put the matter strongly because I am

back on the rigid routine.) The functional normal form provides the bookkeeping device: a complex existential "variable", in effect, namely 'y_x', which will automatically instantiate identically under identical instantiations of the preceding quantifier '$\forall x$' and differently under different instantiations of '$\forall x$'.

Such, then, without talk of functions, is the role of the subscripts on the existential variables. It becomes clear, despite the specificity of the example above, that the method of functional normal forms is just a quick way of doing the work that is done by the main method under the rigid routine.

There is a variant method of functional normal forms, due to Dreben, that does not depend on prenex form. Schemata need little preparation. Conditionals and biconditionals need to be expanded into alternation, conjunction, and negation if their components contain quantifiers, but not otherwise. Negation signs need to be driven inward by DeMorgan's laws and the rules of passage of negation ((9) and (10) of Chapter 23), but only until no quantifier appears in a negated schema. In short, what is required is that quantifiers be overlain only by conjunction, alternation, and quantifiers. Also, as in the previous method of functional normal forms, the existential quantifiers must be relettered so that their variables are distinct from the free variables and from one another.

Thus suppose we want to prove these premises inconsistent:

(1) $\exists x\ \forall y\ Fxy,$

(2) $-\ \exists x \forall z\ \exists y (Fzy\ .\ Fyx),$

(3) $\forall x \forall y [Fxy \rightarrow .\ \forall z (Fzy \rightarrow Fxz) \rightarrow \exists z (Fyz\ .\ Fzx)].$

We prepare (2) by using the rules of passage to drive the negation sign inward:

$$\forall x\ \exists z\ \forall y\ -(Fzy\ .\ Fyx).$$

We prepare (3) by translating the outer two conditional signs and then driving a resulting negation sign inward.

$$\forall x \forall y [-Fxy \vee \exists z\ -(Fzy \rightarrow Fxz) \vee \exists z (Fyz\ .\ Fzx)].$$

We prepare (1)–(3) further by relettering the existential quantifiers for diversity. For clarity I carry this relettering a little farther than is strictly necessary.

$$\exists u\ \forall y\ Fuy,$$
$$\forall x\ \exists v\ \forall y\ -(Fvy\ .\ Fyx),$$
$$\forall x\ \forall y[-Fxy \vee \exists w\ -(Fwy \rightarrow Fxw) \vee \exists t(Fyt\ .\ Ftx)].$$

The reader may have begun to recognize that these are equivalent to the premises (1)–(3) of Chapter 30, which were used again in the present chapter to illustrate the first method of functional normal forms. But let us go ahead now with the new method. We form the functional normal forms again, just as before.

(4) $\forall y\ Fuy,$
(5) $\forall x\ \forall y\ -(Fv_xy\ .\ Fyx),$
(6) $\forall x\ \forall y[-Fxy \vee -(Fw_{xy}y \rightarrow Fxw_{xy}) \vee . Fyt_{xy} . Ft_{xy}x].$

This time the lexicon comprises 'u', 'v_u', 'w_{uu}', 't_{uu}', 'v_{v_u}', and so on. Inconsistent instances are forthcoming.

This bypassing of the prenex form has an effect on the rule for forming the functional normal form, though it did not happen to show itself in the above example. The subscripts on an existential variable should list just the variables of the universal quantifiers whose scopes contained that existential quantifier; and these need not be all the preceding universal quantifiers, when the schema is not prenex. E.g., the schema:

(7) $$\forall x\ \exists y\ Fxy\ .\ \forall z\ \exists w\ Gzw$$

has the functional normal form '$\forall x\ Fxy_x\ .\ \forall z\ Gzw_z$'; the occurrence of '$\forall x$' off to the left does not impose a subscript 'x' on the 'w', because the scope of '$\forall x$' stops short of there.

If we had begun by turning (7) into prenex form so as to subject it to the previous method of functional normal forms, we should have had to settle for one or other of:

$$\forall x\,\exists y\forall z\,\exists w(Fxy\ .\ Gzw), \qquad \forall z\,\exists w\forall x\,\exists y(Fxy\ .\ Gzw),$$

not to mention less prudent choices such as:

$$\forall x\forall z\,\exists y\,\exists w(Fxy\ .\ Gzw).$$

This would mean ending up with one or other of the functional normal forms:

$$\forall x \forall z (Fxy_x \cdot Gzw_{xz}), \qquad \forall z \forall x (Fxy_{zx} \cdot Gzw_z).$$

Either way we get a double subscript, as we did not when we proceeded directly from (7) in Dreben's way. This brings out an advantage of Dreben's method additional to the advantage of sparing us the excursion into prenex form: it sometimes issues in simpler subscripts, thereby simplifying the eventual search for inconsistent instances.

Justification of this method will be omitted here.

EXERCISES

1. Find inconsistent instances of (4)–(6).
2. By way of further exercises in generous number, repeat the examples and exercises of Chapters 29 and 31 by the methods of functional normal forms.

36
HERBRAND'S METHOD

The main method and the methods of functional normal forms were methods of proving inconsistency. Now validity is truth under all interpretations, as inconsistency is falsity under all interpretations. Validity and inconsistency are related as truth and falsity, hence by duality (Chapter 12). Alternation is similarly related to conjunction, and existential quantification to universal. Each of these inconsistency methods, then, can be converted to a validity method by duality. Thus take again the main method. The formal moves undergo little change. UI and EI switch their roles; however, since both were present and both remain present, this only means transferring to UI the requirement that the instantial variable be new. But the goal changes radically; what is wanted now of an accumulation of unquantified instances is not that their con-

junction be truth-functionally inconsistent, but that their alternation be truth-functionally valid. What we prove by achieving this goal, moreover, is not that the conjunction of the premises was inconsistent, but that their alternation was valid.

Thus consider, to begin with, a trivial inconsistency proof by the main method. The top two lines are premises and the succeeding four are instances.

$$\begin{array}{l} \exists x\ \forall y\ Fxy \\ \exists y\ \forall x\ -Fxy \\ \forall y\ Fzy \\ \forall x\ -Fxw \\ Fzw \\ -Fzw \end{array}$$

This proves the inconsistency of the conjunction:

$$\exists x\ \forall y\ Fxy\ .\ \exists y\ \forall x\ -Fxy.$$

Now the validity of the alternation:

$$\forall x\ \exists y\ Fxy \vee \forall y\ \exists x\ -Fxy$$

has the following proof, the dual of the inconsistency proof above:

$$\begin{array}{l} \forall x\ \exists y\ Fxy \\ \forall y\ \exists x\ -Fxy \\ \exists y\ Fzy \\ \exists x\ -Fxw \\ Fzw \\ -Fzw \end{array}$$

It is the validity of the alternation '$Fzw \vee -Fzw$' of these last two instances, not the inconsistency of their conjunction, that attests to the validity of the alternation of the two top schemata.

This same latter proof might be thought to double as a proof of the inconsistency of the conjunction of its top schemata by the main method. This would be wrong; the final step of EI violates the inconsistency method in using an old instantial variable 'z'. Moreover the top schemata are in fact consistent; interpret 'Fxy' as '$x = y$' and they both become true.

From the point of view of implication, the inconsistency proofs and the validity proofs face opposite ways. In the inconsistency proofs, implication tended downward. UI was implicative. EI was not, and that was what complicated the soundness argument. In the validity proofs, implication tends upward. EI is what is implicative here, but upward. The reason the last step of the validity proof above does no harm, despite the staleness of the 'z', is simply that '$-Fzw$' implies '$\exists x\ -Fxw$'. If we were independently to establish the soundness of this proof procedure for validity—which we could do by precisely arguing the soundness proof in Chapter 30 of the main method, duality for duality—we would find that what complicates this argument is that the UI of this method is not implicative upward. And now it is UI, not EI, that has to be hedged by the fresh-variable requirement in order that the justification be made to go through.

For a more substantial example let us form the dual of the inconsistency proof (1)–(12) of Chapter 30. As an aid to writing the dual we think of the '$\rightarrow$' in that proof as '$\vee$' with the antecedent negated. The duals of the old premises will be called "alternative theses," since they are now schemata whose alternation is to be proved valid.

ALTERNATIVE THESES:	(1)	$\forall u\ \exists y\ Fuy$	
	(2)	$\exists x\ \forall v\ \exists y\ -(Fvy \vee Fyx)$	
	(3)	$\exists x \exists y \forall w(-Fxy \cdot -Fxw \vee Fwy \cdot Fyw \vee Fwx)$	
INSTANCES:	(4)	$\exists y\ Fzy$	[of (1)]
	(5)	$\forall v\ \exists y\ -(Fvy \vee Fyz)$	[of (2)]
	(6)	$\exists y\ -(Fty \vee Fyz)$	[of (5)]
	(7)	Fzt	[of (4)]
	(8)	$\exists y \forall w(-Fzy \cdot -Fzw \vee Fwy \cdot Fyw \vee Fwz)$	[of (3)]
	(9)	$\forall w(-Fzt \cdot -Fzw \vee Fwt \cdot Ftw \vee Fwz)$	[of (8)]
	(10)	$-Fzt \cdot -Fzs \vee Fst \cdot Fts \vee Fsz$	[of (9)]
	(11)	Fzs [of (4)]	
	(12)	$-(Fts \vee Fsz)$ [of (6)]	

The alternation of (7), (10), (11), and (12) is found by truth-value analysis to be valid, and this is meant to prove that the alternation of (1), (2), and (3) is valid.

Present methods aside, proofs of validity are intuitively preferable to proofs of inconsistency. Truth is our game, and direct aim is more straightforward than snaring and trapping. However, we see from the examples above that the proof of inconsistency proceeds more naturally under the main method than that of validity under the present method. There are two reasons for this. One is that a downward-trending implica-

tion, in which the implying schema is given and the implied one is sought, comes more easily than the opposite. The other reason is that we find it much more natural to view an accumulating lot of instances conjunctively than in alternation.

This difference in naturalness on the part of the two methods can be partly offset by adjusting the notation. We can turn our earlier little validity proof upside down and render the alternations explicit, *dilemma* fashion:

$$\begin{gathered} Fzw \vee \neg Fzw \\ \exists y\, Fzy \vee \exists x\, \neg Fxw \\ \forall x\, \exists y\, Fxy \vee \forall y\, \exists x\, \neg Fxy \end{gathered}$$

This is indeed natural. We set down, to begin with, an unquantified schema that is palpably truth-functionally valid. Then we derive further lines by operating on the clauses of the alternation. Now that the proof is turned over, these operations are the inverses of UI and EI; hence UG and EG, which is to say *universal* and *existential generalization*. Instead of producing instances from their generics (Chapter 30), they lead from instances to their generics. More specifically they are *clausal* UG and EG, operating as they do on clauses of the new alternations and not just on the whole new lines.

Let us try a similar inversion and transformation of the proof (1)–(12). Our top line now will be the valid alternation of (7), (10), (11), and (12). The next line will be the alternation rather of (4), (10), (4), and (6); thus the clauses (7), (11), and (12) of the first line give way here to their existential generics. We delete the extra (4). The third line will be the alternation of (4), (9), and (6); thus the clause (10) of the second line gives way here to its universal generic (9). The fourth line will be the alternation of (4), (8), and (6); the fifth line will be the alternation of (4), (3), and (5); and the last line will be the alternation of (1), (3), and (2). This is the alternation whose validity was to be proved. The reader is advised to write out the above proof explicitly in full schemata, using wide paper.(*)

This procedure starts with a truth-functionally valid line and derives further valid lines progressively by clausal UG, clausal EG, and, we have seen, a trivial third operation: deletion of duplicate clauses. This third operation accomplishes what, before inversion, had been accomplished by using one line as source of more than one instance. The change in style is caused by our having joined up the instances in alternation.

The proof procedure to which we are led by these rearrangements is one that dates from Herbrand. It exhibits an axiomatic style comparable to what we saw in Chapter 13. Our axiom schemata now, unlike those others, are infinite in number: all the truth-functionally valid schemata qualify outright. This is all right; infinitude of axioms is manageable as long as we have a way of deciding whether a schema is an axiom, and this we have in truth-value analysis. Our rules of inference now are three in number; no longer modus ponens and substitution, but rather clausal UG, clausal EG, and deletion of duplicates.

The two proofs by which we have thus far illustrated the Herbrand method have been needlessly compact, what with the simultaneous generalization of coordinate clauses. The axiomatic structure of the procedure emerges better if we take the steps one at a time. The proof of:

$$\forall x\ \exists y\ Fxy \vee \forall y\ \exists x\ -Fxy,$$

which was represented in three lines above, would then appear rather thus:

$$\begin{aligned} & Fzw \vee -Fzw \\ \exists y\ & Fzy \vee -Fzw \\ \exists y\ & Fzy \vee \exists x\ -Fxw \\ \forall x\ \exists y\ & Fxy \vee \exists x\ -Fxw \\ \forall x\ \exists y\ & Fxy \vee \forall y\ \exists x\ -Fxy \end{aligned}$$

Each line proceeds from its predecessor by a step of clausal UG or EG. The reader would do well to write out the proof of the alternation of (1), (2), and (3) now in this style of one step at a time. There will be eleven long lines altogether, beginning with the truth-functionally valid axiom consisting of (7), (10), (11), and (12). Each succeeding line will proceed from its predecessor by UG or EG or—in one case—deletion of a duplicate. (†)

Let us formulate clausal UG explicitly. We look back to the UI which it inverts. That was UI in the form suited to the style of validity proof with which this chapter began. It was subject to the requirement that the instantial variable be new—i.e., free nowhere hitherto. Now that we have inverted the proof, this newness requirement gives way to an obsolescence requirement: the instantial variable must not still be free anywhere in the derived line.

Indeed, since we now join the old short lines into long ones by alternation, we must not let the instantial variable be free even in the other clauses of alternation within its own line. However, this requirement, that the variable not be free in a coordinate clause, turns out to be assured already by the requirement that it not be free in the derived line. (It is assured as long as we set out our proofs explicitly in the style of one step at a time.) For, if the variable were free in a coordinate clause, then it would be free in the next line too, since those clauses come down.

Clausal UG, then, formulated directly, is just this: *if a line is an alternation of one or more clauses, supplant any of them by a universal quantification whereof it is an instance, provided that the instantial variable is not free in the line thus derived.* It must be stressed that the clauses concerned are strictly clauses of alternation.

It may seem that, symmetrically with the original requirement that the instantial variable be new, we should now require that it be free nowhere in any future lines. In fact this precaution is idle, as long as the variable is not free in the immediately derived line. For our three rules of inference provide no way of introducing a free variable, or of freeing a variable.

Clausal EG is simpler: *if a line is an alternation of one or more clauses, supplant any of them by an existential quantification whereof it is an instance.* There is no restriction on the instantial variable here, as there was none on that of EI in the style of validity proof at the beginning of the chapter.

What is required of a rule of inference in an axiomatic system of valid schemata is not that it be implicative, but just that it transmit validity: that it lead from valid schemata to none but valid schemata. Implicativity is one way of meeting this requirement; schemata implied by valid schemata are indeed valid. But a rule of inference can transmit validity without being implicative. An example is substitution. It is not implicative; '$p \vee qr$' yields '$q \vee \bar{r}s$' by substitution, but does not imply it. What matters is just that when we start with a valid schema, e.g. '$p \vee \bar{p} \vee q$', substitution gives only valid schemata.

Now the systems in Chapter 13 had one rule of inference that was implicative, namely modus ponens, and one that was not, namely substitution. The axiom system newly arrived at has two rules of inference that are implicative and one that is not. Obviously the rule of deletion of duplicates is implicative. Also clausal EG is implicative. For, to begin with, the instance clause implies its generalization; 'Fy' implies '$\exists x\, Fx$'. Further, if a clause A implies a clause B then the alternation of A with any

further material C implies the alternation of B with C; this is evident on truth-functional grounds.

Clausal UG is not implicative, but it transmits validity. This is essentially just the fact, familiar from Chapter 25, that universal quantifications of valid schemata are valid. Further considerations are a relettering and a rule of passage, (4) of Chapter 23. Let us sort out these considerations.

If to a valid schema we apply a universal quantifier governing the whole line, the result is again valid; that is the familiar fact from Chapter 25. Our present UG is only clausal, but it results in a line that is equivalent to the line that could be got by quantifying the whole; this we see from the rule of passage (4) of Chapter 23. Thus look back to the five-line validity proof last displayed. The fourth line comes from the third by clausal UG. The reasoning behind the transmission of validity is that, having found the third line valid, we know that its quantification:

$$\forall z(\exists y\, Fzy \vee \exists x\, {-}Fxw)$$

is valid; and this reduces, by the rule of passage, to:

$$\forall z\ \exists y\, Fzy \vee \exists x\, {-}Fxw,$$

which is the same as the fourth line except for an inconsequential relettering. The step of clausal UG that leads from this fourth line to the last line admits of similar analysis.

The appeal to the rule of passage requires that the variable that we are generalizing not be free in other clauses of the alternation. But we saw that this assurance is provided indirectly by the proviso regarding instantial variables that was incorporated into the formulation of clausal UG.

What now of the inconsequential relettering noted above? Clausal UG allows it; but how does the formulation of clausal UG keep the lettering inconsequential? To reletter 'z' in '$\forall z\ Gzw$' as 'w', e.g., would give '$\forall w\ Gww$' and thus change not just the notation but the structure. It is prevented as follows. What in our present analysis of clausal UG we are picturing as a step first of generalization from 'Gzw' to '$\forall z\ Gzw$', and then of (faulty) relettering to '$\forall w\ Gww$', would have been seen in terms of clausal UG rather as a single step changing the clause 'Gzw' to '$\forall w\ Gww$'; and this step violates clausal UG simply because 'Gzw' is not an instance of '$\forall w\ Gww$'. Granted, 'Gww' is an instance of '$\forall z\ Gzw$'; but not 'Gzw' of '$\forall w\ Gww$'.

Might we then use clausal UG to supplant '*Gww*' somewhere by '$\forall z$ *Gzw*'? No, not that either—for a different reason. The instantial variable, '*w*' here, would continue free in the '$\forall z$ *Gzw*' of the derived line, in violation of the proviso on instantial variables in the formulation of clausal UG.

First, in this chapter, we got a proof procedure for validity from our main method for inconsistency, by duality considerations. Its soundness was provable by an argument dual to that for the soundness of the main method. Then we gave this validity method an axiomatic form, indeed Herbrand's, by inverting it and writing in the alternation signs. Finally, for good measure, we have just seen a direct proof of the soundness of the method thus reformulated. Just as the main method was a method of proving the joint inconsistency of one or more prenex schemata, so this one is a method of proving the validity of an alternation of one or more prenex schemata.

To simplify the formulation and justification of this method I switched to the style, a few pages back, of one step at a time. This imposes the labor, needless in practice, of rewriting the whole long line for each change of a single clause. In practice we would retreat of course to our intermediate condensed style, that of operating on clauses side by side.

It seems, even so, that inconsistency proofs are swifter and easier still—despite the seeming perversity of orientation. There we can leave the premises short and numerous, and the instances likewise, and we can consider them in groups without straining the imagination, all because the relevance of the grouping is conjunctional rather than alternational.

EXERCISES

1. See (∗) and (†). Turn your paper so as to allow long lines.
2. Do examples and exercises from Chapter 29 by Herbrand's method. How is this method a direct extension of that of Chapter 29?

37
OTHER METHODS FOR VALIDITY

We saw how to move from the main method to a validity method by duality, and thence to Herbrand's by inverting the proofs and linking up its lines in alternation. Similar transformations can be performed on the method of functional normal forms (Chapter 35).

Let us begin by simply taking the dual of that method. Universal and existential quantification switch roles. Thus, in forming the functional normal form for validity proofs, we drop universal quantifiers rather than existential ones; and to the recurrences of their variables we attach subscripts recording the variables of prior existential quantifiers. In lieu of premises now we have alternative theses; and their alternation is proved valid by instantiating the existential quantifications by means of lexicon until instances have been accumulated whose alternation is truth-functionally valid.

We may illustrate this method by proving, again, the validity of the alternation of (1)–(3) of the preceding chapter. The proof is simply the dual of the inconsistency proof seen early in Chapter 35, and begins with the new functional normal forms.

ALTERNATIVE THESES:
(1) $\exists y\, Fuy$
(2) $\exists x\ \exists y\ -(Fv_x y \vee Fyx)$
(3) $\exists x\ \exists y(-Fxy \cdot -Fxw_{xy} \vee Fw_{xy}y \cdot Fyw_{xy} \vee Fw_{xy}x)$

INSTANCES:
(4) Fuv_u
(5) $-Fuv_u \cdot -Fuw_{uv_u} \vee Fw_{uv_u}v_u \cdot Fv_u w_{uv_u} \vee Fw_{uv_u}u$
(6) Fuw_{uv_u}
(7) $-(Fv_u w_{uv_u} \vee Fw_{uv_u}u)$

The alternation of (4)–(7) can be seen to be truth-functionally valid.

If we prefer to invert this validity proof and make the alternations explicit, we can make two long lines suffice. Our top line becomes the alternation of (4)–(7), generously spaced so as to facilitate our writing doubly generalized clauses underneath. Then our second line is the alter-

nation of (1), (3), (1) again, and (2); each of these clauses stands beneath the instance that it generalizes. Then we cross out the repetition of (1). What remains, the alternation of (1)–(3), amounts to a functional normal form of the alternation of quantificational schemata whose validity was to be proved (viz. that of (1)–(3) of Chapter 36). (*)

Simultaneous multiple EG has enabled us to get this all at once. This advantage of functional normal forms in validity proofs corresponds to the advantage of them that we noted in inconsistency proofs (Chapter 35), viz. that we could plunge directly to unquantified instances by multiple UI.

Let us look next to Dreben's method of functional normal forms without prenexing. This time I shall go directly to the final method which comes of that one by dualizing, inverting, and marking alternations. We want, say, to prove the validity of:

$$\forall y(\forall x[\exists z(Fxz \,.\, Fzx) \rightarrow -Fxy] \rightarrow \exists x[-Fxy \,.\, \forall w(Fxw \rightarrow -Fwx)]).$$

We prepare it by clearing it of the conditional signs that overlie the quantifiers.

$$\forall y(-\forall x[-\exists z(Fxz \,.\, Fzx) \vee -Fxy] \vee \exists x[-Fxy \,.\, \forall w(Fxw \rightarrow -Fwx)]).$$

Also we drive in all such negation signs.

$$\forall y(\exists x[\exists z(Fxz \,.\, Fzx) \,.\, Fxy] \vee \exists x[-Fxy \,.\, \forall w(Fxw \rightarrow -Fwx)]).$$

We would reletter the universal quantifiers (no longer the existential) if they showed duplications; but they do not. So we move to the functional normal form.

$$(1) \qquad \exists x[\exists z(Fxz \,.\, Fzx) \,.\, Fxy] \vee \exists x(-Fxy \,.\, Fxw_x \rightarrow -Fw_x x).$$

We proceed with the proof by writing down this truth-functionally valid formula.

$$Fyy \,.\, Fyy \,.\, Fyy \,.\vee.\, Fw_y y \,.\, Fyw_y \,.\, Fw_y y \,.\vee.\, -Fyy \,.\, Fyw_y \rightarrow -Fw_y y.$$

Under this, finally, we perform five simultaneous steps of EG. Under the clause '$Fyy \;.\; Fyy \;.\; Fyy$' we write, as its generalization, the first half of (1); under the clause '$Fw_y y \;.\; Fyw_y \;.\; Fw_y y$' we write again the first half of (1); and under the final '$-Fyy \;.\; Fyw_y \rightarrow -Fw_y y$' we write the second half of (1). The resulting line, with its duplicate clause crossed out, is (1) as desired. (†)

Before leaving the topic of proof procedures for validity, we should pause a moment over the axiomatic approach. The Herbrand method in the preceding chapter was seen to fall under this head, with its infinitude of axioms (the truth-functionally valid schemata) and its three rules of inference (clausal UG, clausal EG, and deletion of duplicates). But there are also axiomatizations more similar in character to the truth-functional systems of Chapter 13: systems with just a few axiom schemata and with modus ponens and substitution as rules of inference. This was once the accepted way of developing the subject. A typical axiomatization of the kind, covering both the truth functions and quantification, has as axioms the schemata (1)–(3) of Chapter 13 and, in addition, these:

$$(2) \qquad \forall x\, Fx \rightarrow Fy, \qquad \forall x(p \rightarrow Fx) \rightarrow . p \rightarrow \forall x\, Fx.$$

Its rules of inference include again modus ponens and substitution, but substitution in the sense of Chapter 28. In addition there is a rule allowing the relettering of variables. Also there is simple, nonclausal UG: the attachment of a universal quantifier to a proved line. The system visibly treats only of negation, the conditional, and universal quantification; however, the coverage is extended as desired by translating the other truth-function signs and the existential quantifiers into this notation in the well-known ways.

I pause only for one small sample of a proof in this system. By substituting in the two axioms last stated, and relettering one bound variable, we have:

$$\forall z\, \forall x(Gz \rightarrow Fx) \rightarrow \forall x(Gy \rightarrow Fx),$$
$$\forall x(Gy \rightarrow Fx) \rightarrow . Gy \rightarrow \forall x\, Fx.$$

By substituting in the axiom (1) of Chapter 13, we have:

$$\forall z\, \forall x(Gz \rightarrow Fx) \rightarrow \forall x(Gy \rightarrow Fx) \;.\rightarrow:.$$
$$\forall x(Gy \rightarrow Fx) \rightarrow . Gy \rightarrow \forall x\, Fx \;:\rightarrow:$$
$$\forall z\, \forall x(Gz \rightarrow Fx) \rightarrow . Gy \rightarrow \forall x\, Fx.$$

From these three results, by two steps of modus ponens, we get:

$$\forall z\ \forall x(Gz \rightarrow Fx) \rightarrow .\ Gy \rightarrow \forall x\ Fx.$$

Proof of interesting schemata by this method is hardly worthwhile, though the system is known to be complete.

As urged in Chapter 13, derivation of theorems is unrewarding in domains admitting of a decision procedure. Hence the convenience of taking all truth-functionally valid schemata outright as axioms, as in Herbrand's method. We can carry this principle yet farther, since we have a decision procedure for monadic quantificational schemata: we can take all monadically valid quantificational schemata as axioms.

By this, as noted early in Chapter 28, I mean more than the valid monadic schemata. I mean also the polyadic schemata that can be got from valid monadic ones by substitution. There is a decision procedure not only for validity of monadic schemata but for monadic validity of polyadic schemata.

For, given any schema, make all possible monadic *superstitutions* on it; i.e., frame all the monadic schemata from which it can be got by substitution. They are soon exhausted, if we ignore purely arbitrary differences having to do with the alphabetical choice of schematic letters. E.g., the only possible superstitutions on '$Fyy \rightarrow \exists x\ Fxy$', ignoring such differences, are '$Gy \rightarrow \exists x\ Gx$', '$Gy \rightarrow Hy$', '$Hy \rightarrow \exists x\ Gx$', and further trivialities like '$p \rightarrow \exists x\ Gx$', '$p \rightarrow q$', '$p$'. Then, in principle, we may test all these monadic schemata and see if one is valid. In practice of course one sees immediately which superstitution to make and test.

Given the completeness of such axiom systems as the one last noted, it is easily seen that we get a complete system of quantification theory if we take all monadically valid schemata as axioms and just modus ponens and simple UG as rules of inference. Here is an example of a proof in it. The schemata:

$$Fxy \rightarrow \exists x\ Fxy,$$
$$\forall y(Fxy \rightarrow \exists x\ Fxy) \rightarrow .\ \forall y\ Fxy \rightarrow \forall y\ \exists x\ Fxy,$$
$$\forall x(\forall y\ Fxy \rightarrow \forall y\ \exists x\ Fxy) \rightarrow .\ \exists x\ \forall y\ Fxy \rightarrow \forall y\ \exists x\ Fxy$$

all prove, by test, to be monadically valid and hence to qualify as axioms. (The reader would do well to elicit the relevant valid monadic schemata here by superstitution. The third one is '$\forall x(Gx \rightarrow p) \rightarrow .\ \exists x\ Gx \rightarrow p$'.) Then from the first of the three by UG we have:

$$\forall y(Fxy \rightarrow \exists x\, Fxy).$$

From this and the second of the three we have by modus ponens:

$$\forall y\, Fxy \rightarrow \forall y\ \exists x\, Fxy$$

and thence by UG:

$$\forall x(\forall y\, Fxy \rightarrow \forall y\ \exists x\, Fxy).$$

From this and the last of the three we have by modus ponens our favorite little essentially polyadic theorem:

$$\exists x\ \forall y\, Fxy \rightarrow \forall y\ \exists x\, Fxy.$$

In an adaptation of the condensed proof notation of Chapter 13 we may write the whole proof thus:

$$\begin{array}{ll} [1] & Fxy \rightarrow \exists x\, Fxy \\ [2] & [\forall y\ 1 \rightarrow .]\ \forall y\, Fxy \rightarrow \forall y\ \exists x\, Fxy \\ & [\forall x\ 2 \rightarrow .]\ \exists x\ \forall y\, Fxy \rightarrow \forall y\ \exists x\, Fxy \end{array}$$

The numeral '2' in the last line stands for the line [2] minus its bracketed matter '$[\forall y\ 1 \rightarrow .]$'. The intention of the notation is that each line, with its initial bracketed matter included, is monadically valid.

HISTORICAL NOTE: As remarked at the end of Chapter 30, the axiomatic style illustrated by (2) was Frege's. The particular axioms (2) were Russell's, 1908, except that he used '$\vee$' twice instead of '$\rightarrow$'. His supporting truth-functional axioms were less neat than the Łukasiewicz axioms chosen above, viz. (1)–(3) of Chapter 13.

EXERCISES

1. Carry out (*) and (†).

2. As a basis for further exercises, the last method of the chapter is recommended: derivation from monadically valid schemata using the condensed proof notation. Do various of the examples and exercises of Chapters 29 and 31 again by this method.

38
DEDUCTION

Not only can the main method be inverted to give a direct proof procedure for validity (Chapter 36); it can also be split and partly inverted to give a direct proof procedure for implication. This plan may best be seen through an example. Consider again the example in Chapter 27 about paintings and critics. The reader proved it in Exercise 2 of Chapter 30 by proving that the premise:

(1) $\exists y \forall x (Fy \,.\, Gx \rightarrow Hxy)$

was inconsistent with what the desired conclusion:

(2) $\forall x \exists y (Gx \rightarrow .\, Fy \,.\, Hxy)$

became under negation, namely:

(3) $\exists x \, \forall y - (Gx \rightarrow .\, Fy \,.\, Hxy).$

He proved the inconsistency of (1) and (3) by deriving '$Fz \,.\, Gw \rightarrow Hwz$' from (1) by EI and UI, and deriving '$-(Gw \rightarrow .\, Fz \,.\, Hwz)$' from (3) by similar but suitably intercalated steps, and then noting the truth-functional inconsistency of these two instances. Now the new plan, accentuating the positive, says to proceed not from (1) and (3) to absurdity but from (1) alone to (2). We proceed to the instance '$Fz \,.\, Gw \rightarrow Hwz$' of (1) as before. But this must truth-functionally imply '$Gw \rightarrow .\, Fz \,.\, Hwz$', if it was going to be truth-functionally inconsistent with

'$-(Gw \rightarrow . Fz \,.\, Hwz)$'. So from '$Fz \,.\, Gw \rightarrow Hwz$' we simply infer '$Gw \rightarrow . Fz \,.\, Hwz$' by a new rule TF, *truth-functional inference;* and from this we climb to the desired (2) by steps of EG and UG, the reverse of the steps of EI and UI which in the old proof led from (3) to '$-(Gw \rightarrow . Fz \,.\, Hwz)$'.

In this way we get a *deduction* of (2) from (1). It runs as follows:

PREMISE:	$\exists y \forall x(Fy \,.\, Gx \rightarrow Hxy)$	
DEDUCTION:	$\forall x(Fz \,.\, Gx \rightarrow Hxz)$	(EI)
	$Fz \,.\, Gw \rightarrow Hwz$	(UI)
	$Gw \rightarrow . Fz \,.\, Hwz$	(TF)
	$\exists y(Gw \rightarrow . Fy \,.\, Hwy)$	(EG)
	$\forall x \exists y(Gx \rightarrow . Fy \,.\, Hxy)$	(UG)

The procedure is essentially the same when, as in the following example, the premise is milked for more than one unquantified instance.

PREMISE:	$\forall w \, \forall x \, -(Fxy \,.\, Fxw \,.\, Fwx)$	
DEDUCTION:	$\forall x \, -(Fxy \,.\, Fxy \,.\, Fyx)$	(UI)
	$-(Fyy \,.\, Fyy \,.\, Fyy)$	(UI)
	$-(Fzy \,.\, Fzy \,.\, Fyz)$	(UI)
	$-Fyy \,.\, -Fyz \vee -Fzy$	(TF)
	$\forall w(-Fyy \,.\, -Fyw \vee -Fwy)$	(UG)
	$\exists x \forall w(-Fxy \,.\, -Fxw \vee -Fwx)$	(EG)

A difference between this deduction and the preceding one is that the line got by TF here is got from the two lines preceding it; it is implied by them jointly.

The same thing happens in the next example, which has two premises.

PREMISES:	$\forall x(Gx \rightarrow . Fx \vee Hx)$	
	$\exists x(Gx \,.\, -Hx)$	
DEDUCTION:	$Gy \,.\, -Hy$	(EI)
	$Gy \rightarrow . Fy \vee Hy$	(UI)
	$Fy \,.\, -Hy$	(TF)
	$\exists x(Fx \,.\, -Hx)$	(EG)

But a complication arises when more than one unquantified instance of the conclusion is called for. Suppose we are given the premise:

(4) $\forall y \exists x(Fyy \,.\, Fyx \,.\vee\, Fxy)$

and want the conclusion '$\forall y\ \exists x\ Fxy$'. Under the main method we would negate the conclusion, convert the negation to '$\exists y\ \forall x\ -Fxy$', and then show this to be inconsistent with (4) by deriving inconsistent instances from '$\exists y\ \forall x\ -Fxy$' and (4) as follows:

$$\begin{array}{l} \forall x\ -Fxz \\ \exists x(Fzz \,.\, Fzx \,.\vee\, Fxz) \\ Fzz \,.\, Fzw \,.\vee\, Fwz \\ -Fzz \\ -Fwz \end{array}$$

But when we try to turn this inconsistency proof into a deduction by our trick of partial inversion, we fail; for '$\exists y\ \forall x\ -Fxy$' contributed not just one unquantified instance but two, '$-Fzz$' and '$-Fwz$'. The truth-functional inconsistency of these two with the instance '$Fzz \,.\, Fzw \,.\vee\, Fwz$' of (4) does not entitle us to claim either 'Fzz' or 'Fwz' as a TF consequence of '$Fzz \,.\, Fzw \,.\vee\, Fwz$'; the most we can claim is their alternation, '$Fzz \vee Fwz$'. From this alternation we can get to our desired conclusion '$\forall y\ \exists x\ Fxy$' only by strengthening our rules: two steps of *clausal* EG will yield the alternation of '$\exists x\ Fxz$' with itself, and then TF will get rid of the duplication and we can go on afterward by UG to '$\forall y\ \exists x\ Fxy$'.

So our bolstered deductive procedure uses these rules: UI, EI, TF, clausal EG, and (for similar needs in other examples) clausal UG. Simple EG and UG are still there, being just the special cases of clausal EG and UG where the clause is taken to be the whole line.

Let us now assemble the deduction of '$\forall y\ \exists x\ Fxy$' from (4).

PREMISE:	$\forall y \exists x(Fyy \,.\, Fyx \,.\vee\, Fxy)$	
DEDUCTION:	$\exists x(Fzz \,.\, Fzx \,.\vee\, Fxz)$	(UI)
	$Fzz \,.\, Fzw \,.\vee\, Fwz$	(EI)
	$Fzz \vee Fwz$	(TF)
	$\exists x\ Fxz \vee Fwz$	(cl. EG)
	$\exists x\ Fxz \vee \exists x\ Fxz$	(cl. EG)
	$\exists x\ Fxz$	(TF)
	$\forall y\ \exists x\ Fxy$	(UG)

If the deductions seen thus far in this chapter are compared with the inconsistency proofs to which they are related by partial inversion, it will be seen that what the EI of the inconsistency proof gives way to under inversion is not EG, but UG. For instance the last step of the deduction above goes from '$\exists x\ Fxz$' to the desired conclusion '$\forall y\ \exists x\ Fxy$' by UG; the corresponding step in the inconsistency proof went from the negation of that desired conclusion, viz. '$\exists y\ \forall x\ -Fxy$', to '$\forall x\ -Fxz$' by EI. Consequently, just as EI in the main method was hedged by requiring a new instantial variable, so both EI and UG must now be hedged in some corresponding way. Indeed, since deduction is a way of proving implication, and yet neither EI nor UG is implicative, some such restriction was bound to be needed.

But the nature of the appropriate restriction is no longer so simple and evident as it was in the main method. We cannot require that the instantial variable of UG be new, for UG does not introduce an instantial variable; it eliminates it. Nor would it be sufficient now to invert matters and require, as in the Herbrand method (Chapter 36), that the instantial variable of UG not reappear afterward; for our deductive procedure uses EI as well as UG, and their instantial variables can collide in pernicious ways that this requirement would not prevent. Thus observe this fallacy:

PREMISE:	$\exists x\ Fx$		
DEDUCTION:	Fy	(EI)	
	$\forall x\ Fx$	(UG)	(wrong)

Certainly '$\exists x\ Fx$' does not imply '$\forall x\ Fx$'.

There are, however, two neat restrictions that prove sufficient: *the instantial variables of EI and UG must be different for each such step, and the instantial variable of each such step must be alphabetically later than all free variables of the generic line of that step.* The soundness proof is in the next chapter. If clausal UG or EG is retained, then note further that its instantial variable must be free in no other clause of that line; however, clausal UG and EG will be superseded presently.

The fallacious little deduction last exhibited is disqualified by the ban on using the same instantial variable 'y' for two steps of the sensitive kinds EI and UG. Actually the big deduction just previously exhibited is disqualified too, in its EI step; for the 'z' in the generic line is alphabetically later than the instantial variable 'w'. However, this violation can be rectified without loss; we can simply use a variable alphabetically later than 'z', say 'z'', instead of 'w' in the three lines where 'w' appeared.

From the readiness of this adjustment we see that the alphabetical restriction is really tighter than necessary. Its virtue is that it is simpler than its more liberal alternatives, and moreover the cases that it needlessly excludes are easily adjusted.

A real fallacy that is excluded by the alphabetical restriction, and cannot be patched up by relettering, is this:

PREMISE:	$\forall x\ \exists y\ Fxy$		
DEDUCTION:	$\exists y\ Fwy$	(UI)	
	Fwz	(EI)	
	$\forall x\ Fxz$	(UG)	(wrong)
	$\exists y\ \forall x\ Fxy$	(EG)	

The free variable 'z' in the line got by UG is alphabetically later than the instantial variable 'w' of that UG step. And if we rectify this by using 'z'' instead of 'w', we thereupon violate the alphabetical restriction in the EI step.

In order to facilitate conformity to these restrictions, and also for a further purpose that will be noted presently, it is convenient to *flag* each step of EI and UG by recording the instantial variable off to the right of the line which that step yields. Also it is convenient, and increasingly so in complex deductions, to number the lines at the left and then refer back to them, as sources, at the right. Indeed the citation of rules (UI, TF, etc.) is less to the point; we recognize the rule easily enough, given the source. Also, in preparation for a certain later departure, let us signal the premise no longer by name but rather by starting a column of stars there and running the column down the length of the deduction. Each star means that the starred line is affirmed only conditionally on the premise.

Here, then, rewritten in the new style and rectified in the alphabetical detail, is the long deduction last previously exhibited.

*(1)	$\forall y\,\exists x(Fyy\ .\ Fyx\ .\vee\ Fxy)$		
*(2)	$\exists x(Fzz\ .\ Fzx\ .\vee\ Fxz)$	(1)	
*(3)	$Fzz\ .\ Fzz'\ .\vee\ Fz'z$	(2)	z'
*(4)	$Fzz \vee Fz'z$	(3)	
*(5)	$\exists x\ Fxz \vee Fz'z$	(4)	
*(6)	$\exists x\ Fxz \vee \exists x\ Fxz$	(5)	
*(7)	$\exists x\ Fxz$	(6)	
*(8)	$\forall y\ \exists x\ Fxy$	(7)	z

The earlier little deduction which used two premises would now appear as follows.

*(1)	$\forall x(Gx \rightarrow . Fx \vee Hx)$	
**(2)	$\exists x(Gx \,.\, -Hx)$	
**(3)	$Gy \,.\, -Hy$	(2) y
**(4)	$Gy \rightarrow . Fy \vee Hy$	(1)
**(5)	$Fy \,.\, -Hy$	(3)(4)
**(6)	$\exists x(Fx \,.\, -Hx)$	(5)

By reference to the flagging device, our restrictions on EI and UG can now be phrased more succinctly: no variable is to be flagged twice, and the flag is to be alphabetically later than all free variables of the generic.

The status claimed for the last line in a deduction, or indeed any line, is that it is implied by its premises. Reservations are needed, however, in connection with flagged variables. This implication claim is in order provided that no flagged variable is free in the line that is said to be implied, nor in its premises. Thus the last deduction above shows that (1) and (2) therein imply (6), but not that they imply (5). And the reader can satisfy himself that they in fact do not. Again, take the deduction before last and imagine line (2) thereof as its premise, with (1) dropped; this deduction would not claim to show that (2) implies (8), since 'z' is free in that premise (2). A deduction is called *finished* if no flagged variable is free in its last line nor in any premise on which that last line depends; and then the point of a finished deduction is that its last line is implied by its premises.

There is a device that has had some currency, in deduction, under the name of *conditional proof*. Its adoption has the effect of dispensing with the clausal forms of UG and EG after all. It has also the more important effect of dispensing wholly, in deductions, with the requirement of prenex form. Also a certain naturalness can be claimed for it. It consists in adopting, temporarily, an additional premise, and eventually discharging it by *conditionalization* (Cd): by incorporating it into the conclusion as antecedent of a conditional. The pattern is this: on the assumption that p, we show that q; so we conclude that $p \rightarrow q$.

The procedure requires some bookkeeping, to keep track of what lines depend on the temporary premise. This is where our notation of stars, rather pointless thus far, comes into its own. We star the temporary premise and each line that depends on it. If yet another temporary premise is adopted while one such is already in force, we begin a further column of stars yet farther to the right.

As an example let us return once more to the deduction before last and see how it might have been managed with help of conditional proof.

*(1)	$\forall y \exists x(Fyy \,.\, Fyx \,.\vee\, Fxy)$		
*(2)	$\exists x(Fzz \,.\, Fzx \,.\vee\, Fxz)$	(1)	
*(3)	$Fzz \,.\, Fzz' \,.\vee\, Fz'z$	(2)	z'
**(4)	Fzz		
**(5)	$\exists x\, Fxz$	(4)	
*(6)	$Fzz \rightarrow \exists x\, Fxz$	*(5)	
**(7)	$Fz'z$		
**(8)	$\exists x\, Fxz$	(7)	
*(9)	$Fz'z \rightarrow \exists x\, Fxz$	*(8)	
*(10)	$\exists x\, Fxz$	(3)(6)(9)	
*(11)	$\forall y\ \exists x\, Fxy$	(10)	z

The column of stars belonging to a premise is finished for good when interrupted. The little right-hand column of two stars at (4) and (5) belongs to the temporary premise 'Fzz' and has nothing to do with the similar column at (7) and (8), which belongs to the temporary premise '$Fz'z$'.

The deduction above could be extended still by adding an unstarred line:

(12) $\forall y \exists x(Fyy \,.\, Fyx \,.\vee\, Fxy) \rightarrow \forall y\ \exists x\, Fxy$ *(11)

thus absorbing the original premise itself into the conclusion. When the last line of a finished deduction is unstarred it claims validity outright; it is no mere consequence of a premise.

A final example illustrates the utility of temporary premises and conditionalization in probing schemata that are not prenex. It is the one again about paintings and critics, which was done at the beginning of the present chapter. Now, however, we render the premise and conclusion directly in the forms that they received in Chapter 27. We neither put them into prenex form nor negate the conclusion:

*(1)	$\exists y[Fy \,.\, \forall x(Gx \rightarrow Hxy)]$		
*(2)	$Fz \,.\, \forall x(Gx \rightarrow Hxz)$	(1)	z
*(3)	$\forall x(Gx \rightarrow Hxz)$	(2)	
*(4)	$Gw \rightarrow Hwz$	(3)	
**(5)	Gw		

**(6)	$Fz \cdot Hwz$	(2)(4)(5)
**(7)	$\exists y(Fy \cdot Hwy)$	(6)
*(8)	$Gw \rightarrow \exists y(Fy \cdot Hwy)$	*(7)
*(9)	$\forall x[Gx \rightarrow \exists y(Fy \cdot Hxy)]$	(8) w

HISTORICAL NOTE: The method set forth in the present pages is of a type known as *natural deduction,* and stems, in its broadest outlines, from Gentzen and Jaśkowski (1934). The rule of conditionalization, which is the crux of natural deduction, appeared as an explicit formal rule somewhat earlier, having been derived by Herbrand (1930) and also in effect by Tarski (1929) from systems of their own of a type other than natural deduction. The derivation consisted in showing that whenever one statement could be deduced from another in the concerned system, the conditional formed of the two statements could also be proved as a theorem by the original rules of that system. In this status of derived rule relative to one system or another, the rule of conditionalization has come to be known in the literature as the *theorem of deduction*.

Jaśkowski's system of natural deduction is conspicuously unlike that of the present pages; for Jaśkowski dispenses with EG and EI, and gets along with milder restrictions on UG, by the expedient of treating '$\exists x$', '$\exists y$', etc., as abbreviations of '$-\forall x-$', '$-\forall y-$', etc. This course is economical in rules, but greatly increases the difficulty and complexity of the deductions themselves. The crucial difference between Gentzen's system and that of the present pages is in EI; he had a more devious rule.

Because of the presence of EI, the present system differs from Gentzen's and Jaśkowski's considerably on the score of restrictions upon the rules. In particular the device of flagging is a novelty. Gentzen and Jaśkowski had restrictions too, but gave them different forms. Cooley, in pp. 126–140 of his *Primer of Formal Logic* (1942), made use of natural deduction in a form which included substantially EI, but without exact formulation of restrictions. Explicitly formulated rules and restrictions, resembling the present system except for wide variations in the restrictions, were set forth by Rosser and independently by me in mimeographed lecture notes from 1946 on.

EXERCISES

1. Check each of the following to see whether it is a correct deduction according to the rules, and a finished one.

*(1)	Fy			*(1)	$\forall x(Fx \to Gx)$		
*(2)	$\forall x\, Fx$	(1)	y	*(2)	$Fy \to Gy$	(1)	
(3)	$Fy \to \forall x\, Fx$	*(2)		*(3)	$\exists x(Fy \to Gx)$	(2)	
				*(4)	$\forall w\, \exists x(Fw \to Gx)$	(3)	y
*(1)	$\forall x(Fx \vee Gy)$			*(1)	$\forall x(Fx \,.\, Gx)$		
*(2)	$Fy \vee Gy$	(1)		*(2)	$Fy \,.\, Gy$	(1)	
*(3)	$\forall x(Fy \vee Gx)$	(2)	y	*(3)	Fy	(2)	
				*(4)	$\forall x\, Fx$	(3)	y
*(1)	$Fx \vee Gy$			*(5)	Gy	(2)	
*(2)	$\forall y(Fx \vee Gy)$	(1)	y	*(6)	$\forall x\, Gx$	(5)	y
*(3)	$\exists x \forall y(Fx \vee Gy)$	(2)		*(7)	$\forall x\, Fx \,.\, \forall x\, Gx$	(4)(6)	

2. If any of the deductions above is correct but unfinished, append further lines so as to produce a finished deduction. If any is incorrect but can be revised into a correct deduction to the same effect, so revise it.

3. Do the examples and various exercises of Chapters 29 and 31 again by deduction.

39
SOUNDNESS

The deductive method last set forth dominated the first two editions of this book. Its rules, we see, take longer to explain than those of what I am now calling the main method, and its soundness, we shall see, takes longer to prove. Once established, the method has its virtues. It requires no prenexing, it is directed on implication rather than inconsistency or even validity, the steps are intuitive on the whole, and the search for proofs tends to go smoothly.

This last is not always true; sometimes an obvious implication, quickly proved by the main method, is stubborn under deduction. For efficiency, on the average, the method of functional normal forms is perhaps best. As for the main method, it is often effortless, as we know; it lends itself to the easiest of completeness proofs; and it opens the way, which we have followed, to a half dozen alternative methods that are easily derived from it and are efficient or illuminating in their several ways. Thus it was that in

the third edition I demoted deduction, limiting my treatment of it to what we saw in the preceding chapter.

The move was deplored by teachers who were accustomed to the early editions and wedded to my method of deduction. Hence I am adding the present omissible chapter and the next, in which I sketch the soundness proof and discuss deductive techniques.

The soundness of the method means that the last line of a finished deduction is implied by its premises (or is valid, if free of premises). To see this we need pause only over Cd, UG, and EI, since UI, EG, and TF are implicative. What needs to be shown in the case of Cd is that if a line in a deduction is implied by its premises, and then we prefix its last premise to it by '$\rightarrow$', the conditional thus formed is implied by the remaining premises. Now the key to this is the equivalence of '$pq \rightarrow r$' to '$p \rightarrow . q \rightarrow r$'. For, think of '$q \rightarrow r$' as representing the inferred conditional and 'p' the remaining premises. The line represented as 'r' was implied by its premises, represented as 'pq'; so '$pq \rightarrow r$' represents a valid conditional. So then does '$p \rightarrow . q \rightarrow r$'. Thus the inferred conditional ('$q \rightarrow r$') is implied by the remaining premises ('p').

It remains then to consider UG and EI. What is clear thus far is that if a deduction lacks UG and EI, in other words if it has no flagged variables, then it is *sound:* its last line is implied by its premises. Now let us assume that finished deductions with just n flagged variables are sound, and on that assumption let us proceed to prove that finished deductions with $n + 1$ flagged variables are likewise sound.

Consider then a finished deduction with $n + 1$ flagged variables of which the alphabetically earliest is 'y'. Suppose it flags a step of UG. Thus the flagged line has the form '$\forall x(\ldots x \ldots)$' and is inferred from a line '$\ldots y \ldots$'. Modify the deduction by prefixing a whole new column of stars and the new premise:

$$*(0) \quad \ldots y \ldots \rightarrow \forall x(\ldots x \ldots).$$

Since the line '$\forall x(\ldots x \ldots)$' is derivable from the lines '$\ldots y \ldots$' and (0) by TF, we can drop the flag 'y'. By the alphabetical rule, 'y' is the alphabetically latest free variable in '$\ldots y \ldots$'; yet it was the alphabetically earliest flagged variable in the deduction; so no flagged variables are free in (0) now that 'y' is unflagged. But neither are any of them free in the last line of the deduction or in its other premises, since the original deduction was a finished one. Therefore the revised deduction is a finished one. But it has only n flagged variables, and we assumed that such deductions are

sound. Therefore its last line is implied by its premises, including (0). So the conditional is valid whose consequent is that last line and whose antecedent is the conjunction of (0) and the other premises. Apply '$\forall y$' to that conditional; the result is still valid. But 'y' is not free in the consequent nor anywhere in the antecedent except in (0), since the original deduction was a finished one. So, by the rules of passage (8) and (1) of Chapter 23, we can change '$\forall x$' to '$\exists x$' and make it govern just (0). The resulting quantification:

$$\exists y[\ldots y\ldots \rightarrow \forall x(\ldots x\ldots)]$$

is valid by (13) of Chapter 28 and so can be dropped from the antecedent, leaving the rest of the conditional still valid. So the last line of the deduction is implied by the old premises.

We assumed that the step using the alphabetically earliest flagged variable was by UG. The argument is parallel for EI. (0) is then '$\exists x(\ldots x\ldots) \rightarrow \ldots y\ldots$', and the pertinent law in Chapter 28 is (14). See how close the reasoning is to the soundness proof in Chapter 30.

We began by observing that deductions without flagged variables are sound. Next, assuming that finished ones with just n flagged variables are sound, I proved that those with $n + 1$ flagged variables are sound. We can now conclude that all finished deductions are sound. This way of reasoning is called *mathematical induction*.

EXERCISE

Wherein does the proof of soundness depend on the requirement that no variable be flagged twice? How do we know that our appeal just now to (13) and (14) complies with the last three lines of page 180?

40 DEDUCTIVE STRATEGY

The following six simple deductions, establishing again three of our familiar rules of passage, will be convenient for reference in illustrating strategies of deduction.

Deduction 1:

*(1)	$\exists x(p \vee Fx)$	
*(2)	$p \vee Fy$	(1) y
**(3)	Fy	
**(4)	$\exists x\, Fx$	(3)
*(5)	$Fy \rightarrow \exists x\, Fx$	*(4)
*(6)	$p \vee \exists x\, Fx$	(2)(5)

Deduction 2:

*(1)	$p \vee \exists x\, Fx$	
**(2)	$\exists x\, Fx$	
**(3)	Fy	(2) y
*(4)	$\exists x\, Fx \rightarrow Fy$	*(3)
*(5)	$p \vee Fy$	(1)(4)
*(6)	$\exists x(p \vee Fx)$	(5)

Deduction 3:

* (1)	$\exists x(p \cdot Fx)$	
* (2)	p	(1) x
* (3)	Fx	
* (4)	$\exists x\, Fx$	(3)
* (5)	$p \cdot \exists x\, Fx$	(2)(4)

Deduction 4:

* (1)	p	
* (2)	$\exists x\, Fx$	
* (3)	Fx	(2) x
* (4)	$p \cdot Fx$	(1)(3)
* (5)	$\exists x(p \cdot Fx)$	(4)

Deduction 5:

*(1)	$\exists x(Fx \rightarrow p)$	
*(2)	$Fx \rightarrow p$	(1) x
**(3)	$\forall x\, Fx$	
**(4)	Fx	(3)
**(5)	p	(2)(4)
*(6)	$\forall x\, Fx \rightarrow p$	*(5)

Deduction 6:

*(1)	$\forall x\, Fx \rightarrow p$	
**(2)	Fx	
**(3)	$\forall x\, Fx$	(2) x
**(4)	p	(1)(3)
*(5)	$Fx \rightarrow p$	*(4)
*(6)	$\exists x(Fx \rightarrow p)$	(5)

The obvious operation of separating conjunctions by TF may conveniently be left tacit by stating conjunctive lines at will as bracketed pairs of lines. This is done in Deductions 3 and 4.

For the discovery of a desired deduction there are some simple strategies. When as in Deduction 1 we are starting with a quantification,

the obvious strategy is to begin by dropping the quantifier by UI or EI (with or without change of variable). Conversely, when as in Deduction 2 we are hoping to get a quantification as end result, the strategy is to try to deduce the desired schema without its quantifier (and with or without change of variable); afterward the quantifier may be supplied by UG or EG. If we were trying to discover Deductions 1 and 2, these two strategies would afford us just this much: (1)–(2) of Deduction 1 and, working backward, (5)–(6) of Deduction 2.

So, if we were trying to devise Deduction 1, the initial strategy of dropping the quantifier would leave us with the problem of getting from (2) of Deduction 1 to (6). Next, since (2) and (6) are alike to the extent of '$p \vee$', it would be natural to wonder whether a conditional joining the remaining parts of (2) and (6), viz., '$Fy \rightarrow \exists x\, Fx$', might combine with (2) to imply (6) truth-functionally. A fell swoop confirms the notion: '$p \vee q$' and '$q \rightarrow r$' in conjunction imply '$p \vee r$'. So we now know that the desired (6) can be subjoined by TF to (2) and (5); the problem remains merely of getting (5). Here an obvious strategy of the conditional comes into play: assume the desired antecedent as additional temporary premise, try to deduce the consequent, and then get the conditional by Cd. So we assume (3), from which (4) happens to proceed without difficulty; and thus Deduction 1 has been created in full.

If we were trying to devise Deduction 2, the strategy of quantifiers would have left us with the problem of getting from (1) of Deduction 2 to (5). The reasoning which solves this problem is exactly parallel to that detailed in the preceding paragraph.

Truth-functional implications are checked easily enough once they are thought of. But when we are building a deduction, the implication has to be thought up before it can be tested. The way in which Deduction 1 was achieved illustrates a course which must often be followed. There we consulted common sense for a suggestion of obtainable lines from which the desired result (6) might follow. (2) was at hand, and common sense suggested (5) as an adequate supplementation. So then we checked the suggestion and found that (2) and (5) do in conjunction imply (6) truth-functionally. Accordingly we undertook to get (5).

The strategy behind the discovery of Deduction 3 is evident. The usual strategy of quantifiers leads from (1) to (2) and (3), leaving us with the problem of getting from (2) and (3) to (5); and this problem offers little challenge to ingenuity. In Deduction 4 the strategy of quantifiers leaves us with the problem of getting (4) from (1) and (2), which again is the work of a moment.

The strategy behind the discovery of Deduction 5 is as follows. The strategy of quantifiers leads us from (1) to (2), leaving us with the problem of getting (6) from (2). According to the strategy of the conditional, in order to get (6) we assume its antecedent as (3) and try to deduce its consequent 'p'. The strategy of quantifiers leads us from (3) to (4), so that all that remains to be done is get 'p' somehow from the lines (1)–(4) which are now at our disposal. Obviously (2) and (4) serve the purpose, via TF.

The strategy behind Deduction 6 is rather as follows. Since we want (6), the strategy of quantifiers directs us to overlook its quantifier and aim for (5). According to the strategy of the conditional, in order to get (5) we assume its antecedent (2) and try to deduce its consequent 'p'. So all that remains to be done is get 'p' somehow from (1) and (2). The intervening line (3) quickly suggests itself.

It must be remembered that our rules of deduction apply only to whole lines. UI and EI serve to remove a quantifier only if the quantifier is initial to a line and covers the line as a whole; and UG and EG serve to introduce a quantifier only into such a position. It would be fallacious, e.g., to proceed to the last line '$\forall x\, Fx \rightarrow p$' of Deduction 5 from the earlier line '$Fx \rightarrow p$' by UG, and it would be fallacious to proceed from the first line '$\forall x\, Fx \rightarrow p$' of Deduction 6 to the subsequent line '$Fx \rightarrow p$' by UI. '$\forall x\, Fx \rightarrow p$' is not a quantification, but a conditional containing a quantification '$\forall x\, Fx$'. What issues from '$Fx \rightarrow p$' by UG, and yields '$Fx \rightarrow p$' by UI, is not '$\forall x\, Fx \rightarrow p$' but '$\forall x(Fx \rightarrow p)$'.

Next let us undertake the inference about slovenly persons at the beginning of Chapter 17. Here our premises and conclusion are:

$$\forall x(Gx \rightarrow .\ Fx \vee Hx), \qquad \exists x(Gx\ .\ -Hx), \qquad \exists x(Fx\ .\ -Hx).$$

The strategy of quantifiers reduces the problem to that of getting from '$Gx \rightarrow .\ Fx \vee Hx$' and '$Gx\ .\ -Hx$' to '$Fx\ .\ -Hx$'. If by luck the conjunction of '$Gx \rightarrow .\ Fx \vee Hx$' and '$Gx\ .\ -Hx$' truth-functionally implies '$Fx\ .\ -Hx$', then our deduction is complete. So we submit:

$$p \rightarrow .\ q \vee r\ :\ p\bar{r}\ :\!\rightarrow q\bar{r}$$

to a truth-value analysis, and find that luck is with us. In full, then,

$$\begin{array}{lll} * & (1) & \forall x(Gx \rightarrow .\ Fx \vee Hx) \\ * & (2) & \exists x(Gx\ .\ -Hx) \\ * & (3) & Gx \rightarrow .\ Fx \vee Hx \qquad (1) \end{array}$$

*	(4)	$Gx \cdot -Hx$	(2) x
*	(5)	$Fx \cdot -Hx$	(3)(4)
*	(6)	$\exists x(Fx \cdot -Hx)$	(5)

A more substantial example is the traditional one about drawing circles. The premise and conclusion were schematized in Chapter 27:

$$\forall x(Fx \rightarrow Gx), \qquad \forall y[\exists x(Fx \cdot Hyx) \rightarrow \exists x(Gx \cdot Hyx)]$$

Now the steps of deduction from the one to the other are dictated almost automatically by the strategies of quantifiers and the conditional. The desired conclusion being a universal quantification, we first aim for this expression minus its '$\forall y$'. But this is a conditional; so we assume its antecedent '$\exists x(Fx \cdot Hyx)$' and try for its consequent '$\exists x(Gx \cdot Hyx)$'. But in order to get '$\exists x(Gx \cdot Hyx)$' the strategy is to try for '$Gx \cdot Hyx$' (or '$Gz \cdot Hyz$', etc.). What we have to deduce this from are '$\forall x(Fx \rightarrow Gx)$' and '$\exists x(Fx \cdot Hyx)$'; so the strategy of dropping quantifiers is brought to bear on these, and little proves to be left to the imagination. In full the deduction is as follows.

*(1)		$\forall x(Fx \rightarrow Gx)$	
**(2)		$\exists x(Fx \cdot Hyx)$	
**(3)		$Fz \cdot Hyz$	(2) z
**(4)		$Fz \rightarrow Gz$	(1)
**(5)		$Gz \cdot Hyz$	(3)(4)
**(6)		$\exists x(Gx \cdot Hyx)$	(5)
*(7)		$\exists x(Fx \cdot Hyx) \rightarrow \exists x(Gx \cdot Hyx)$	*(6)
*(8)		$\forall y[\exists x(Fx \cdot Hyx) \rightarrow \exists x(Gx \cdot Hyx)]$	(7) y

Note that the shift from 'x' to 'z', in line (3), was necessitated by the alphabetical stipulation in EI. (We could omit most such shifts by letting the order of the alphabet vary from deduction to deduction.)

Next there is the example about paintings and critics in Chapter 27. We already went through the deduction twice in Chapter 38. All the steps of deduction are dictated by the strategies of quantifiers and the conditional. Let us record the deduction anew and review its genesis.

*	(1)	$\exists y[Fy \cdot \forall x(Gx \rightarrow Hxy)]$	
*	(2)	Fy	(1) y
*	(3)	$\forall x(Gx \rightarrow Hxy)$	(1) y
*	(4)	$Gx \rightarrow Hxy$	(3)

**(5)	Gx	
**(6)	$Fy \,.\, Hxy$	(2)(4)(5)
**(7)	$\exists y(Fy \,.\, Hxy)$	(6)
*(8)	$Gx \rightarrow \exists y(Fy \,.\, Hxy)$	*(7)
*(9)	$\forall x[Gx \rightarrow \exists y(Fy \,.\, Hxy)]$	(8) x

Lines (2)–(4) issue automatically from our strategy of dropping quantifiers. Moreover, since we want (9), the backward strategy of quantifiers directs us to aim for (8); and in order to get (8) the strategy of the conditional directs us to assume (5) and aim for (7). In order to get (7) we try for (6), according to the backward strategy of quantifiers. So now the deduction is complete if in fact (6) happens to follow from its predecessors by TF. Truth-value analysis or inspection shows that it does. Thus the truth-functional implication leading to (6) from (2), (4), and (5) did not need to be thought up; it automatically presented itself for appraisal.

In the example of Chapter 27 about philosophers, the premise and conclusion have the respective forms:

$$\exists y[Fy \,.\, \forall x(Fx \rightarrow Gxy)], \qquad \exists x(Fx \,.\, Gxx).$$

Our strategies of quantifiers prompt us to derive 'Fy' and '$\forall x(Fx \rightarrow Gxy)$' from the premise, and to aim for 'Fx' and 'Gxx'—or say 'Fy' and 'Gyy'. So all that remains is to get 'Gyy' from 'Fy' and '$\forall x(Fx \rightarrow Gxy)$', which is easy. The full deduction, then, is this:

* (1)	$\exists y[Fy \,.\, \forall x(Fx \rightarrow Gxy)]$		
* (2)	Fy	} (1)	y
* (3)	$\forall x(Fx \rightarrow Gxy)$	}	
* (4)	$Fy \rightarrow Gyy$	(3)	
* (5)	$Fy \,.\, Gyy$	(2)(4)	
* (6)	$\exists x(Fx \,.\, Gxx)$	(5)	

Note that the changes of variables between (3) and (4) and between (5) and (6) are not, as in previous examples, prompted by the restrictions attending UG and EI. They are simply steps in the obvious route from (2) and (3) to (6).

As an example of a deduction calling for more ingenuity in the manipulation of variables, let us show that symmetry and transitivity together imply reflexivity. Our problem is to deduce '$\forall x \forall y(Fxy \rightarrow . \, Fxx \,.$

Fyy)' from the symmetry premise '$\forall x \forall y(Fxy \rightarrow Fyx)$' and the transitivity premise '$\forall x \forall y \forall z(Fxy \,.\, Fyz \mathrel{.\!\!\rightarrow} Fxz)$'. Backward strategy tells us to aim for '$Fxy \rightarrow\!. \; Fxx \,.\, Fyy$', or perhaps '$Fuw \rightarrow\!. \; Fuu \,.\, Fww$', afterward getting the desired conclusion by two applications of UG. In deductions which promise to involve serious shuffling of variables, we can often obviate awkward conflicts of variables by immediately shifting to wholly new variables in our backward strategy; so let us aim for '$Fuw \rightarrow\!. \; Fuu \,.\, Fww$'. To get this, the strategy is to assume 'Fuw' and try for both 'Fuu' and 'Fww'. So the problem reduces to that of getting 'Fuu' and 'Fww' from 'Fuw' and the two original premises of symmetry and transitivity. Turning then to forward strategy, we consider dropping the quantifiers from the premises; but there remains the problem of picking new variables in suitable ways. We are well advised to pick them as 'u' and 'w' exclusively, since only these appear in the desired results 'Fuu' and 'Fww' and the intermediate premise 'Fuw'. In the symmetry premise the relettering '$Fuw \rightarrow Fwu$' is more promising than '$Fwu \rightarrow Fuw$', since our intermediate premise 'Fuw' will combine with '$Fuw \rightarrow Fwu$' to yield something more, 'Fwu', with which to work. So now we have 'Fuw' *and* 'Fwu' to go on. Therefore the two reletterings of the transitivity premise which we seem to have to choose between, viz., '$Fuw \,.\, Fwu \mathrel{.\!\!\rightarrow} Fuu$' and '$Fwu \,.\, Fuw \mathrel{.\!\!\rightarrow} Fww$', will both be useful; one will yield our desired result 'Fuu', and the other will yield our other desired result 'Fww'. So our deduction uses the transitivity premise twice, and runs as follows (subject to subsequent refinement):

* (1)	$\forall x \forall y(Fxy \rightarrow Fyx)$	
* (2)	$\forall x \forall y \forall z(Fxy \,.\, Fyz \mathrel{.\!\!\rightarrow} Fxz)$	
* (3)	$\forall y(Fuy \rightarrow Fyu)$	(1)
* (4)	$Fuw \rightarrow Fwu$	(3)
* (5)	$\forall y \forall z(Fuy \,.\, Fyz \mathrel{.\!\!\rightarrow} Fuz)$	(2)
* (6)	$\forall z(Fuw \,.\, Fwz \mathrel{.\!\!\rightarrow} Fuz)$	(5)
* (7)	$Fuw \,.\, Fwu \mathrel{.\!\!\rightarrow} Fuu$	(6)
* (8)	$\forall y \forall z(Fwy \,.\, Fyz \mathrel{.\!\!\rightarrow} Fwz)$	(2)
* (9)	$\forall z(Fwu \,.\, Fuz \mathrel{.\!\!\rightarrow} Fwz)$	(8)
*(10)	$Fwu \,.\, Fuw \mathrel{.\!\!\rightarrow} Fww$	(9)
**(11)	Fuw	
**(12)	Fwu	(11)(4)
**(13)	Fuu	(11)(12)(7)
	Fww	(12)(11)(10)

*(14)	$Fuw \rightarrow . Fuu \,.\, Fww$	*(13)
*(15)	$\forall y(Fuy \rightarrow . Fuu \,.\, Fyy)$	(14) w
*(16)	$\forall x \forall y(Fxy \rightarrow . Fxx \,.\, Fyy)$	(15) u

Once a deduction is discovered, it is easily enough revised so as to eliminate unnecessary intervening steps. A stretch of unquantified deduction such as appears in lines (4), (7), (10)–(14) is bound to embody a single truth-functional implication. In the process of discovery we built up the implication piecemeal, but now that its end points are visible we can verify mechanically that (14) is truth-functionally implied directly by the conjunction of (4), (7), and (10). So in retrospect we can refine our deduction by deleting (11)–(13) and justifying (14) directly by the citation '(4)(7)(10)'. Incidentally, for condensation we might omit (8)–(9) and just write 'similarly' after (10).

The advantage of having aimed for (14) in the form '$Fuw \rightarrow . Fuu \,.\, Fww$', rather than in the form '$Fxy \rightarrow . Fxx \,.\, Fyy$', may be appreciated by rewriting the above deduction with 'x' and 'y' in place of 'u' and 'w' everywhere. Difficulty will be found to arise in (8).

The strategy of reductio ad absurdum, which underlay the main method, can usefully be revived where other strategies fail. It consists in assuming the contradictory of what is to be proved and then looking for trouble. It may be illustrated by deducing '$- \exists x\, Fx$' from '$\forall x\, -Fx$':

*(1)	$\forall x\, -Fx$	
*(2)	$-Fx$	(1)
**(3)	$\exists x\, Fx$	
**(4)	Fx	(3) x
*(5)	$\exists x\, Fx \rightarrow Fx$	*(4)
*(6)	$- \exists x\, Fx$	(2)(5)

Here the usual strategy of quantifiers leads from (1) to (2), leaving us with the problem of proceeding thence to (6). Resorting to reductio ad absurdum, we assume the contradictory of the desired (6) as (3). Thence by the strategy of quantifiers we move to (4), and find the trouble we were looking for; for (4) conflicts with (2). Cd and TF then lead through (5) to (6).

A supplementary strategy worth remarking is that of the *dilemma,* which is useful in getting a conclusion from an alternation. First deduce the desired conclusion separately from each component of the alternation, and derive a conditional in each case by Cd; then from these conditionals and the original alternation infer the conclusion by TF. See page 243.

EXERCISES

1. In strict analogy to the deduction last presented, deduce '$- \forall x\, Fx$' from '$\exists x\, -Fx$'.

2. Establish the rest of the rules of passage, namely (2), (4), (5), (6), and (8) of Chapter 23, by pairs of deductions.

3. Establish by deduction the syllogisms in Chapter 16, including the reinforced one about Spartans. Model: the above deduction about slovenly persons.

4. Similarly for the inference about witnesses, middle of Chapter 17.

5. Deduce '$\exists x\, -Fx$' from '$- \forall x\, Fx$'. Plan on having this intermediate line:

$$*(4) \quad Fx \rightarrow \forall x\, Fx \quad *(3)$$

Analogously, deduce '$\forall x\, -Fx$' from '$- \exists x\, Fx$'.

6. Establish the equivalence of '$\forall x(p \,.\, Fx \rightarrow Fx)$' to '$p$' by mutual deduction. Hint: 'p' truth-functionally implies '$p \,.\, Fx \rightarrow Fx$'.

7. Establish the equivalence of '$\exists x(p \,.\, Fx \rightarrow Fx)$' to '$p$'.

8. The following are excerpts from a deductive solution of Exercise 4 of Chapter 25. Complete it.

$*(3)$	$Gx \rightarrow Hx$	(1)
$**(4)$	$\exists x(Fx \,.\, Gx)$	
$**(5)$	$Fx \,.\, Gx$	(4) x
$***(7)$	$\forall x(Hx \rightarrow -Jx)$	(6) y
$**(10)$	$Hy \rightarrow -Jy \,.\!\rightarrow.\, Fx \rightarrow -Hx$	$*(9)$
$**(11)$	$Hy \,.\, Jy$	(3)(5)(10)

9. Establish the inference about the class of '00, end of Chapter 17. Warning: This deduction is a holiday venture. None of the truth-functional implications involved is as formidable as that which leads to (11) in the preceding exercise, but the deduction runs to 18 lines (in my version anyway), some of which are not easily come by. It tends to heighten one's appreciation of the solution in Chapter 19.

10. Deduce 'Fyy' from '$\forall x\, Fxy \vee \forall x\, Fyx$', following the strategy of the dilemma.

IV
GLIMPSES BEYOND

41
SINGULAR TERMS

The logic of truth functions and quantification is now under control. Some simple sorts of inference, however, still want discussion—notably those turning upon *singular terms* such as 'Socrates':

All men are mortal,	Socrates is a Greek,
Socrates is a man;	Socrates is wise;
∴ Socrates is mortal.	∴ Some Greeks are wise.

Furthermore the theory of identity, including such evident laws as '$x = x$' and '$x = y \mathrel{.\leftrightarrow.} y = x$', remains untouched. A few chapters will suffice to do justice to singular terms and identity. There will be an optional excursion into something called term functors, and in conclusion we shall have a brief glimpse of set theory, or the theory of classes—a discipline intimately linked to logic but better regarded as the basic discipline of classical mathematics.

What were called "terms" in Chapters 14 and 27 and represented by 'F', 'G', etc., are *general terms,* as opposed to singular terms. But generality is not to be confused with ambiguity. The singular term 'Jones' is ambiguous in that it might be used in different contexts to name any of various persons, but it is still a singular term in that it purports in any particular context to name one and only one person. The same is true even of pronouns such as 'I' and 'thou'; these again are singular terms, but merely happen to be highly ambiguous pending determination through the context or other circumstances attending any given use of them. The same may be said of 'the man', or more clearly 'the President', 'the cellar'; these phrases (unlike 'man', 'president', and 'cellar' themselves) are singular terms, but the one and only one object to which they purport to refer in any given use depends on attendant circumstances for its determination.

Besides the classification of terms into singular and general, there is a cross classification into *concrete* and *abstract*. Concrete terms are those which purport to refer to individuals, physical objects, events; abstract terms are those which purport to refer to abstract objects, e.g., to numbers, classes, attributes. Thus some singular terms, e.g., 'Socrates',

'Cerberus', 'earth', 'the author of *Waverley*', are concrete, while other singular terms, e.g., '7', '3 + 4', 'piety', are abstract. Again some general terms, e.g., 'man', 'house', 'red house', are concrete (since each man or house is a concrete individual), while others, e.g., 'prime number', 'zoological species', 'virtue', are abstract (since each number is itself an abstract object, if anything, and similarly for each species and each virtue).

Caution then comes to be needed in our talk of term abstracts, to which we have been accustoming ourselves; for a term abstract may or may not be an abstract term. A term abstract is a general term, and an abstract one only if the several objects of which it is true happen to be abstract. In Chapter 46 we shall indeed contemplate the bold step of letting general terms do double duty as names of their extensions and thus letting term abstracts serve as abstract singular terms, but only on a par with other general terms.

The division of terms into concrete and abstract is a distinction only in the kinds of objects referred to. The distinction between singular and general terms is more vital from a logical point of view. Thus far I have drawn it only in a vague way: a term is singular if it purports to name an object (one and only one), and otherwise general. Note the key word 'purports'; it separates the question off from such questions of fact as the existence of Socrates and Cerberus. Whether a word purports to name one and only one object is a question of language, and is not contingent on facts of existence.

In terms of logical structure, what it means to say that the singular term "purports to name one and only one object" is just this: *The singular term belongs in positions of the kind in which it would also be coherent to use variables* '*x*', '*y*', *etc.* (or, in ordinary language, pronouns). Contexts like:

Socrates is wise, Piety is a virtue,
Cerberus guards the gate, $7 = 3 + 4$,

etc., are parallel in form to open sentences:

x is wise, x guards the gate, x is a virtue, $x = 3 + 4$

such as may occur in closed statements having the form of quantifications: '$\exists x(x$ is wise)', etc. The terms 'Socrates', 'Cerberus', 'piety', and '7'

are, in short, substitutable for variables in open sentences without violence to grammar; and it is this that makes them singular terms. Whether there is in fact such an object as Socrates (which, tenselessly, there is) or Cerberus (which there is not) or piety or 7 (on which philosophers disagree) is of course a separate question.

General terms, in contrast to singular ones, do not occur in positions appropriate to variables. Typical positions of the general term 'man' are seen in 'Socrates is a man', 'All men are mortal'; it would not make sense to write:

(1) Socrates is an x, All x are mortal,

or to imbed such expressions in quantifications in the fashion:

(2) $\exists x$(Socrates is an x),
(3) $\forall x$(all x are mortal $\rightarrow$ Socrates is mortal).

The 'x' of an open sentence may refer to objects of any kind, but it is supposed to refer to them one at a time; and then application of '$\forall x$' or '$\exists x$' means that what the open sentence says of x is true of all or some objects taken thus one at a time.

There are indeed legitimate open sentences somewhat resembling (1) but phrased in terms of class membership, thus:

(4) Socrates is a member of x, All members of x are mortal.

But these do not, like (1), show 'x' in place of a general term such as 'man'; rather they show 'x' in place of an abstract singular term, 'mankind' ('class of all men'), as in 'Socrates is a member of mankind', 'All members of mankind are mortal'. The open sentences (4) may quite properly appear in quantifications:

(5) $\exists x$(Socrates is a member of x),
(6) $\forall x$(all members of x are mortal $\rightarrow$ Socrates is mortal).

Incidentally (6) can be further analyzed:

(7) $\forall x[\forall y(y$ is a member of $x \rightarrow y$ is mortal) $\rightarrow$ Socrates is mortal].

As an alternative to (4) we might also appeal to attributes instead of classes, thus:

(8) Socrates has x, Everything that has x is mortal.

Here 'x' appears in the position of an abstract singular term such as 'humanity' which purports to name an attribute. Quantifications analogous to (5)–(7) can then be built on (8).

Just as the sentence letters in a schema stand as dummy sentences and the term letters as dummy general terms, so the free variables may be seen as standing as dummy singular terms. To represent the above syllogisms about Socrates schematically, then, we may simply use a free 'y' to represent 'Socrates'. The inferences are then justified directly by the two valid conditionals:

$$\forall x(Fx \rightarrow Gx) \,.\, Fy \mathrel{.\!\!\rightarrow} Gy, \qquad Fy \,.\, Gy \mathrel{.\!\!\rightarrow} \exists x(Fx \,.\, Gx).$$

Another example:

PREMISES: Aldrich bribed every member of the committee,
Barr is a member of the committee;
CONCLUSION: Someone bribed Barr.

This is justified by the validity of the conditional:

$$\forall x(Fx \rightarrow Gzx) \,.\, Fw \mathrel{.\!\!\rightarrow} \exists x\, Gxw,$$

which goes into a pure existential and is quickly checked.

These valid conditionals show that the conclusions will come out true if the premises do, no matter what objects of the universe we choose in interpretation of 'y', 'z', and 'w'. In particular therefore we may choose Socrates, Aldrich, and Barr—provided merely that the universe contains such things.

But this last proviso is essential to the intended application of our deductive results. A singular term may or may not name an object. A singular term always *purports* to name an object, but is powerless to guarantee that the alleged object be forthcoming; witness 'Cerberus', The deductive techniques of quantification theory with free variables serve very well for inferences depending on singular terms when we are assured that *there are* objects such as those terms purport to name; so this question

of existence then becomes the central question where singular terms are concerned.

I shall find no use for the narrow sense which some philosophers have given to 'existence', as against 'being'; viz., concreteness in space-time. If any such special connotation threatens in the present pages, imagine 'exists' replaced by 'is'. When the Parthenon and the number 7 are said to be, no distinction in the sense of 'be' need be intended. The Parthenon is indeed a placed and dated object in space-time while the number 7 (if such there be) is another sort of thing; but this is a difference between the objects concerned and not between senses of 'be'.

In contrast to 7 and the Parthenon, there is no such thing as Cerberus; and there is no such number as 0/0. Clearly these repudiations do not of themselves depend on any limitation of existence to space-time. The meaning of the particular word 'Cerberus' merely happens to be such that, if the word did name an object, that object would be a physical object in space and time. The word 'Cerberus' is like 'Parthenon' and 'Bucephalus' in this respect, and unlike '7' and '0/0'. But the word 'Cerberus' differs from 'Parthenon' and 'Bucephalus' in that whereas there is something in space-time such as the word 'Parthenon' purports to name (viz., at Athens for some dozens of centuries including part or all of the twentieth), and whereas there is (tenselessly) presumably something in space and time such as the word 'Bucephalus' purports to name (viz., at a succession of positions in the Near and Middle East in the fourth century B.C.), on the other hand there happens to be nothing such as the word 'Cerberus' purports to name, near or remote, past, present, or future.

It is surely a commonplace that some singular terms may, though purporting to name, flatly fail to name anything at all. 'Cerberus' is one example, and '0/0' is another. But, commonplace though this be, experience shows that recognition of it is beset with persistent confusions, to the detriment of a clear understanding of the logic of singular terms. Let us make it our business in the remainder of this chapter to dispel certain of these confusions.

There is a tendency to try to preserve some shadowy entity under the word 'Cerberus', for example, lest the word lose its meaning. If 'Cerberus' were meaningless, not only would poetry suffer, but even certain blunt statements of fact, such as that there is no such thing as Cerberus, would lapse into meaninglessness. Thus we may hear it said, e.g., that Cerberus exists as an idea in the mind. But this verbal maneuver conduces only to confusion. Of a tangible object such as the Parthenon, to change

the subject for a moment, it would be wanton obscurantism to affirm a *double* existence: in Athens *and* in the mind. Far more straightforward to admit two (or many) objects: the tangible Parthenon in Athens, and the Parthenon-idea in the mind (or the Parthenon-ideas in many minds). 'Parthenon' names the Parthenon and only the Parthenon, whereas 'the Parthenon-idea' names the Parthenon-idea. Similarly not 'Cerberus', but 'the Cerberus-idea', names the Cerberus-idea; whereas 'Cerberus', as it happens, names nothing.

This is not the place to try to say what an idea is, or what existence in the mind means. Perhaps from the point of view of experimental psychology an idea should be explained somehow as a propensity to certain patterns of reaction to words or other stimuli of specified kinds; and perhaps "existing in the mind" then means simply "being an idea." But no matter; the idea of "idea" is entertained here only as a concession to the other party. The point is that though we be as liberal about countenancing ideas and other nonphysical objects as anyone may ask, still to identify the Parthenon with the Parthenon-idea is simply to confuse one thing with another; and to try to assure there being such a thing as Cerberus by identifying it with the Cerberus-idea is to make a similar confusion.

The effort to preserve meaning for 'Cerberus' by presenting some shadowy entity for 'Cerberus' to name is misdirected; 'Cerberus' remains meaningful despite not naming. Most words, like 'and' or 'sake', are quite meaningful without even purporting to be names at all. Even when a word is a name of something, its meaning would appear not to be identifiable with the thing named.[1] Mount Everest has been known, from opposite points of view, both as Everest and as Chomolungma;[2] here the named object was always one, yet the names can scarcely be viewed as having been alike in *meaning* or synonymous; for no insight into the combined minds of all users of 'Everest' and 'Chomolungma' could reveal that these named the same thing, pending a strenuous investigation of nature. Again there is Frege's example of 'Evening Star' and 'Morning Star'; the named planet is one, but it took astronomy and not mere analysis of meanings to establish the fact.

Precise and satisfactory formulation of the notion of meaning is an unsolved problem of semantics. Perhaps the meaning of a word is best construed as the associated idea, in some sense of 'idea' which needs to

[1] This much neglected point was well urged by Frege, "Über Sinn und Bedeutung."
[2] Erwin Schrödinger, *What is Life?*, last paragraph, nearly enough.

be made precise in turn; or perhaps as the system of implicit rules in conformity with which the word is used, supposing that a criterion of "implicit rule" can be devised which is selective enough to allow sameness of meaning on the part of distinct expressions. Perhaps, indeed, the best treatment of the matter will prove to consist in abandoning all notion of so-called meanings as entities; thus such phrases as 'having *meaning*' and 'same *in meaning*' might be dropped in favor of 'significant' and 'synonymous', in hopes eventually of devising adequate criteria of significance and synonymy involving no excursion through a realm of intermediary entities called meanings. Perhaps it will even be found that of these only significance admits of a satisfactory criterion, and that all effort to make sense of 'synonymy' must be abandoned along with the notion of meaning.[3] However all this may be, the important point for present purposes is that significance of a word, even of a word which (like 'Cerberus') purports to be a name, is in no way contingent upon its naming anything; and even if a word does name an object, and even if we countenance entities called meanings, there is still no call for the named object to be the meaning.

The mistaken view that the word 'Cerberus' must name something in order to mean anything turns, it has just now been suggested, on confusion of naming with meaning. But the view is encouraged also by another factor, viz., our habit of thinking in terms of the misleading word 'about'. If there is no such thing as Cerberus, then, it is asked, what are you talking *about* when you use the word 'Cerberus' (even to say that there is no such thing)? Actually this protest could be made with the same cogency (viz., none) in countless cases where no would-be name such as 'Cerberus' occurs at all: What are you talking about when you say that there are no Bolivian battleships? The remedy here is simply to give up the unwarranted notion that talking sense always necessitates there being things talked about. The notion springs, no doubt, from essentially the same confusion which was just previously railed against; then it was confusion between meanings and objects named, and now it is confusion between meanings and things talked about.

This mistaken view that 'Cerberus' must name something has been seen to evoke, as one lame effort to supply a named object, the notion that Cerberus is something in the mind. Other expedients to the same end are commonly encountered. There is, e.g., the relativistic doctrine according to which Cerberus exists in the world of Greek mythology and not in the

[3] See my *From a Logical Point of View*, Essays II and III.

world of modern science. This is a perverse way of saying merely that Greeks believed Cerberus to exist and that (if we may trust modern science thus far) they were wrong. Myths which affirm the existence of Cerberus have esthetic value and anthropological significance; moreover they have internal structures upon which our regular logical techniques can be brought to bear; but it does happen that the myths are literally false, and it is sheer obscurantism to phrase the matter otherwise. There is really only one world, and there is not, never was, and never will be any such thing as Cerberus.

Another such expedient which had best not detain us long, if only because of the mazes of metaphysical controversy in which it would involve us if we were to tarry, is the view that concrete individuals are of two kinds: those which are actualized and those which are possible but not actualized. Cerberus is of the latter kind, according to this view; so that there *is* such a thing as Cerberus, and the proper content of the vulgar denial of Cerberus is more correctly expressed in the fashion 'Cerberus is not actualized'. The universe in the broader sense becomes badly overpopulated under this view, but the comfort of it is supposed to be that there comes to *be* something Cerberus about which we may be said to be talking when we rightly say (in lieu of 'There is no such thing as Cerberus') 'Cerberus is not actualized'.

But this device depends on the possibility of Cerberus; it no longer applies when we shift our example from 'Cerberus' to some would-be name of complex form involving an out-and-out impossibility, e.g., 'the cylindrical pyramid of Copilco'. Having already cluttered the universe with an implausible lot of unactualized possibles, are we to go on and add a realm of unactualizable impossibles? The tendency at this point is to choose the other horn of the supposed dilemma, and rule that expressions involving impossibility are meaningless. Thence, perhaps, the not uncommon notion that a statement logically inconsistent in form must be reclassified as a nonstatement: not false but meaningless. This notion, besides being unnatural on the face of it, is impractical in that it rules out the possibility of tests of meaningfulness; for logical consistency admits of no general test, nor even of a complete proof procedure (see Chapter 34).

All this piling of expedients on expedients is, insofar as prompted by a notion that expressions must name to be meaningful, quite uncalled for. There need be no mystery about attributing nonexistence where there is nothing to attribute it to, and there need be no misgivings over the mean-

ingfulness of words which purport to name and fail. To purport to name and fail is already proof of a full share of meaning.

If for *other* reasons the recognition of unactualized possibles is felt to be desirable, there is nothing in the ensuing logical theory that need conflict with it as long as essential distinctions are preserved. So-called mental ideas and so-called meanings were provisionally tolerated above; "ideas" of the Platonic stripe, including unactualized possibles, could be accommodated as well. What must be insisted on is merely that such shadowy entities, if admitted, be named in some distinguishing fashion: 'the Cerberus-possibility', or 'the Cerberus-idea', or 'the meaning of 'Cerberus' '. If we can get together on this much by way of convention, then everyone can be left to his favorite metaphysics so far as anything further is concerned. The essential message to be carried over from this chapter into succeeding ones is simply this: Some meaningful words which are proper names from a grammatical point of view, notably 'Cerberus', do not name anything.[4]

EXERCISES

1. Show that the premises:

   Barr and Williams did not both contribute,
   If Blake contributed then so did everybody

   imply the conclusion:

   Blake did not contribute,

   allowing free variables to stand for the proper names.

2. Show that the premises:

   Edith envies everyone more fortunate than she,
   Herbert is no more fortunate than any who envy him

   imply the conclusion:

   Herbert is no more fortunate than Edith.

[4] For more on this theme see Russell, "On denoting," and my *From a Logical Point of View*, Essays I and VI.

42
IDENTITY

Identity is such a simple and fundamental idea that it is hard to explain otherwise than through mere synonyms. To say that x and y are identical is to say that they are the same thing. Everything is identical with itself and with nothing else. But despite its simplicity, identity invites confusion. E.g., it may be asked: Of what use is the notion of identity if identifying an object with itself is trivial and identifying it with anything else is false?

This particular confusion is cleared up by reflecting that there are really not just two kinds of cases to consider, one trivial and the other false, but three:

$$\text{Cicero} = \text{Cicero}, \qquad \text{Cicero} = \text{Catiline}, \qquad \text{Cicero} = \text{Tully}.$$

The first of these is trivial and the second false, but the third is neither trivial nor false. The third is informative, because it joins two different terms; and at the same time it is true, because the two terms are names of the same object. For truth of a statement of identity it is necessary only that '=' appear between names of the same object; the names may, and in useful cases will, themselves be different. For it is not the names that are affirmed to be identical, it is the things named. Cicero is identical with Tully (same man), even though the name 'Cicero' is different from the name 'Tully'. To say anything about given objects we adjoin appropriate words to *names* of the objects; but there is no reason to expect that what is thereby said of the objects will be true also of the names themselves. The Nile, e.g., is longer than the Tuscaloosahatchie, but the names are oppositely related.

Since the useful statements of identity are those in which the named objects are the same and the names are different, it is only because of a peculiarity of language that the notion of identity is needed.[5] Still, no linguistic investigation of the names in a statement of identity will suffice, ordinarily, to determine whether the identity holds or fails. The identities:

[5] Thus it was that Hume had trouble accounting for the origin of the identity idea in experience. See *Treatise of Human Nature*, Bk. I, Pt. IV, Sec. II.

Everest = Chomolungma (see Chapter 41),
Evening Star = Morning Star,
25th President of U.S. = first President of U.S. inaugurated at 42,
Mean temperature at Tuxtla = 93°F

all depend for their substantiation upon inquiry into extra-linguistic matters of fact.

A popular riddle, commonly associated with identity, is this: How can a thing that changes its substance be said to remain identical with itself? How is it, e.g., that one's body may be spoken of as the same body over a period of years? The problem dates from Heraclitus, who said "You cannot step into the same river twice, for fresh waters are ever flowing in upon you." Actually the key to this difficulty is to be sought not in the idea of identity but in the ideas of thing and time. A physical thing—whether a river or a human body or a stone—is at any one moment a sum of simultaneous momentary stages of spatially scattered atoms or other small physical constituents. Now just as the thing at a moment is a sum of these spatially small parts, so we may think of the thing over a period as a sum of the temporally small parts which are its successive momentary stages. Combining these conceptions, we see the thing as extended in time and in space alike; the thing becomes a sum of momentary stages of particles, or briefly particle-moments, scattered over a stretch of time as well as space. All this applies as well to the river or human body as to the stone. There is only a difference of detail in the two cases: in the case of the stone the constituent particle-moments pair off fairly completely from one date to another as momentary stages of the same particles, whereas in the case of the river or human body there is more heterogeneity in this respect. The river or human body will regularly contain some momentary stages of a particle and exclude other momentary stages of the same particle, whereas with the stone, barring small peripheral changes or ultimate destruction, this is not the case. Here we have a distinction reminiscent of the distinction in traditional philosophy between "modes" and "substances". But things of both kinds are physical things in one and the same sense: sums of particle-moments. And each thing is identical with itself; we *can* step into the same river twice. What we cannot do is step into the same temporal part of the river twice, where the part is temporally shorter than a stepping-while. Diversity among the parts of a whole must not be allowed to obscure the identity of the whole, nor of each part, with itself.

Thus far we have been thinking of statements of identity composed of '=' flanked by singular terms. But '=' is an ordinary relative term, and so may be flanked as well by variables; e.g.:

$\forall x \forall y(x \text{ is a god} \,.\, y \text{ is a god} \,.\!\rightarrow\!.\, x = y)$,
$\exists x[x \text{ is a god} \,.\, \forall y(y \text{ is a god} \rightarrow .\, x = y)]$,
$\exists x \exists y(x \text{ is a god} \,.\, y \text{ is a god} \,.\, x \neq y)$,
$\forall x \forall y \forall z(x \text{ is a god} \,.\, y \text{ is a god} \,.\, z \text{ is a god} \,.\!\rightarrow\!: x = y \,.\vee.\, x = z \,.\vee.\, y = z)$.

(The notation '$x \neq y$' is a convenient abbreviation for '$-(x = y)$'.) As the reader can verify on a little reflection, these four statements amount respectively to the following:

There is one god at most,[6]
There is exactly one god,
There are at least two gods,
There are two gods at most.

Statements of identity consisting of '=' flanked by singular terms are needed because two singular terms may name one thing. But the need of '=' flanked by variables arises from a different peculiarity of language; viz., from its use of multiple variables of quantification (or their pronominal analogues in ordinary language). Two variables are allowed to refer to the same object, and they are also allowed to refer to different objects; and thus the sign of identity comes to be needed when, as in the above four examples, there arises the question of sameness or difference of reference on the part of the variables. From a logical point of view it is the use of the identity sign between variables, rather than between singular terms, that is fundamental. We shall see, indeed (Chapter 44), that the whole category of singular terms is theoretically superfluous, and that there are logical advantages in thinking of it as theoretically cleared away.

The logic of identity is a branch not reducible to the logic of quantification. Its notation may be thought of as comprising that of the logic of quantification plus the one additional sign '='. Thus the schemata of the logic of identity are the same as the quantificational schemata except that they may contain, along with the clauses 'p', 'q', 'Fx', 'Gxy', etc., additional clauses of the form of identities: '$x = y$', '$x = z$', etc. Validity may be defined for schemata of the logic of identity precisely as it was

[6] This, according to a quip of Whitehead's, is the creed of the Unitarians.

defined for quantificational schemata (Chapter 27). Thus there are valid schemata such as:

$$\exists y(x = y \,.\, Fy) \rightarrow \exists y(x = y)$$

which are like valid quantificational schemata except that '$x = x$', '$x = y$', etc., turn up instead of schematic clauses 'Gxx', 'Gxy', etc. These are valid still by virtue simply of their quantificational structure, and independently of any peculiarities of identity. But there are further valid schemata whose validity depends specifically on the meaning of identity; one such is:

(I) $$Fx \,.\, x = y \,.\!\rightarrow Fy.$$

For, consider any choice of universe, any interpretation of 'F' therein, and any assignment of objects to the free variables 'x' and 'y'. If the object assigned to 'x' is the same as that assigned to 'y', and is an object of which 'F' is interpreted as true, then (I) comes out true through truth of its consequent 'Fy'; and in any other case (I) comes out true through falsity of its antecedent.

Through having added the identity sign to our logical notation we find ourselves able, for the first time, to write genuine sentences without straying from our logical notation. Hitherto, schemata were the best we could get; extralogical materials had to be imported from ordinary language when, for the sake of illustration, a genuine sentence was wanted. In '$x = x$' and '$\forall x(x = x)$', however, we have sentences—the first open, the second closed and true.

Any schemata obtained by putting identities '$x = x$', '$x = y$', etc., in place of say 'Gxx', 'Gxy', etc., in valid quantificational schemata are, it has been explained, valid schemata of the logic of identity. But such putting of identities for schematic clauses can, when thorough, yield a sentence instead of a schema, e.g.:

$$x = x \,.\!\rightarrow \exists y(x = y).$$

The concept of validity may be applied to such sentences along with the schemata, following the precedent of Chapter 28.

This example is even quantificationally valid, in the sense of that chapter. Of those whose validity depends specifically on the meaning of identity, on the other hand, the simplest is:

(II) $$x = x.$$

Another, truth-functionally implied by the two foregoing sentences, is '$\exists y(x = y)$'.

Substitutions may be made for 'F' in (I) just as explained in Chapter 28, except that the substituted expressions may now contain identity signs instead of term letters. Reflection on the general mechanics of substitution (Chapter 26) reveals that, in (I) in particular, 'Fx' and 'Fy' may in effect be directly supplanted respectively by any schemata that differ only in that the one has free 'x' in some places where the other has free 'y'. For, picture the two schemata as '$\ldots x \ldots x \ldots y \ldots$' and '$\ldots x \ldots y \ldots y \ldots$'; then the abstract theoretically substituted for 'F' would be '$\{z\colon \ldots x \ldots z \ldots y \ldots\}$'. Thus one result of substitution in (I) is:

$$w = x \,.\, x = y \,.\!\!\rightarrow\!. \, w = y$$

—the law of transitivity of identity (cf. Chapter 29).

The universal closures of (I) and (II), viz:

$$\forall x \forall y(Fx \,.\, x = y \,.\!\!\rightarrow\, Fy), \qquad \forall x(x = x),$$

together with the universal closures of any results of substituting for 'F' in (I), will be called *axioms of identity*. Now the technique of proof in quantification theory carries over to the logic of identity; we simply treat the axioms as premises. E.g., the law '$x = y \,.\!\!\rightarrow\!. \, y = x$' of the symmetry of identity follows truth-functionally from the case:

$$x = x \,.\, x = y \,.\!\!\rightarrow\!. \, y = x$$

of (I) together with (II).

For further illustration let us show '$\forall y(x = y \,.\!\!\rightarrow\, Fy)$' equivalent to '$Fx$', by establishing implication forward and backward in the style of the end of Chapter 37.

$$[\text{II} \rightarrow.] \; \forall y(x = y \,.\!\!\rightarrow\, Fy) \rightarrow Fx.$$
$$[\forall y \; \text{I} \rightarrow.] \; Fx \rightarrow \forall y(x = y \,.\!\!\rightarrow\, Fy).$$

Each of these full lines is monadically valid.

It must be remembered that in these proofs the identity sign behaves merely as an inert predicate letter, as if '$x = x$' and '$x = y$' were 'Gxx'

and '*Gxy*'. The peculiar properties of identity are invoked explicitly, when wanted, by citing an axiom of identity. It is convenient sometimes to cite the axiom with its quantifier or quantifiers and sometimes without, as the above examples show.

HISTORICAL NOTE: Gödel showed in 1930 (Theorem VII) that these axioms of identity are complete. That is, every valid schema or sentence of the logic of identity can be proved from them by the logic of quantification.

EXERCISES

1. Writing '*y*' for 'Barr', '*z*' for 'the cashier', and '*F*' for 'had a key', put the statement:

   None but Barr and the cashier had a key

   into logical notation with help of identity. Show that this and:

   Someone who had a key took the briefcase

   imply the conclusion:

   Barr or the cashier took the briefcase

   with the help of an axiom of identity.

2. Prove the equivalence of:

   $$\exists x(Fx \,.\, x = y), \qquad Fy$$

   by proving the two implications with help of axioms of identity.

43
DESCRIPTIONS

It is usual in logic to write '$(\imath x)$', with an inverted iota, to mean 'the object x such that'. Thus the complex singular terms 'the author of *Waverley*' and 'the prime number between 5 and 11' become:

$$(\imath x)(x \text{ wrote } \textit{Waverley}), \qquad (\imath x)(x \text{ is prime } . \ 5 < x < 11).$$

Singular terms are called *descriptions* when written in this form. The singular terms of ordinary language which may be represented thus as descriptions begin typically with the singular 'the', but by no means necessarily so, as these examples show:

what he went after,	$(\imath x)$(he went after x);
where he was born,	$(\imath x)$(he was born at x);
John's mother,	$(\imath x)(x$ bore John);
Smith's house,	$(\imath x)(x$ is a house . x is Smith's).

In general a singular term purports to name one and only one object, and in particular a singular term of the form '$(\imath x)Fx$' purports to name the one and only object of which the general term represented by 'F' is true. Thus, if y is the object $(\imath x)Fx$, then y must be such that

$$Fy \ . \ F \text{ nothing-but-}y.$$

This conjunction amounts to saying that, for each thing x, 'F' is true of x if $x = y$, and false of x otherwise. In short:

(1) $$\forall x(Fx \leftrightarrow . \ x = y).$$

E.g., to say that Scott is $(\imath x)(x$ wrote *Waverley*$)$ is to say that

$$\forall x(x \text{ wrote } \textit{Waverley} \leftrightarrow . \ x = \text{Scott}).$$

If 'F' is true of nothing or of many things, then there is no such thing as $(\imath x)Fx$. Actually the term appearing in the role of the 'F' of '$(\imath x)Fx$' in

verbal examples from ordinary discourse very frequently needs supplementary clauses to narrow it down to the point of being true of only one object, but this situation can commonly be viewed merely as a case of the familiar practice of depending on context or situation to resolve ambiguities of ordinary language. Moreover, it would be unnatural to construe all use of the singular 'the' in this way; often a better account is simply pronominal. Commonly 'the boy', 'the car', serves merely as a pronoun whose grammatical antecedent is some name or description or perhaps a general term and quantifier. In logical notation such a pronoun might come through simply as a bound variable.

We saw in Chapter 41 that arguments involving a singular term can be carried through by straight quantification theory with a free variable, say 'y', for the singular term, but that the application of the results depends on construing y as the object named by the singular term, and hence is contingent on existence of such an object. This construing of y, and the existence assumption on which it rests, figured nowhere in the schematism of the proof, but only in the informal step of application. Now the beauty of descriptions is that here the construing of y as the named object can itself be schematized quite explicitly as an additional premise, of the form (1) above. So our technique for arguments involving descriptions is as follows: we use free variables for the descriptions as for any singular terms, but we also add a *descriptional premise* of the form (1) for each description. An axiom of identity is also usually called for, because of the '=' in the descriptional premise. Thus let us try this example:

PREMISE: The broker who hired John hired only honors graduates,
CONCLUSION: John was an honors graduate.

Here we have two singular terms, the simple one 'John' and the complex one 'the broker who hired John'. Let us represent them by free variables 'w' and 'y' respectively. Writing 'F' for 'broker' and 'G' for 'hired', we may also render the complex singular term as a description '$(\imath x)(Fx \,.\, Gxw)$'; so the corresponding descriptional premise is:

$$\forall x(Fx \,.\, Gxw \,.\leftrightarrow.\, x = y).$$

So from this and the original premise, which is '$\forall x(Gyx \rightarrow Hx)$' where '$H$' represents 'honors graduate', and perhaps an axiom of identity in

addition, we want to show that '*Hw*' follows. We do so by transforming the conditional:

$$\forall x(Fx \,.\, Gxw \,.\leftrightarrow.\, x = y) \,.\, \forall x(Gyx \rightarrow Hx) \,.\, \forall x(x = x) \,.\rightarrow\, Hw$$

into a pure existential and finding it valid.

Whether a proposed deduction is to enjoy the benefits of a descriptional premise depends, evidently, on whether a given singular term can fairly be translated into the form of a description. Now fairness of translation is a vague matter, hinging as it does on the concept of synonymy which was so dimly regarded in Chapter 41. 'The author of *Waverley*' seems fairly translatable as '$(\imath x)$(x wrote *Waverley*)', but 'Scott' and 'the author of *Ivanhoe*' do not, despite the fact that all of these name the same object; for it is felt that 'the author of *Waverley*' is connected with '$(\imath x)$(x wrote *Waverley*)' by sheer meaning, whereas 'Scott' and 'the author of *Ivanhoe*' are connected with it through accidental matters of fact.

At the same time it seems that singular terms can depart widely in form from the singular 'the' idiom and still be fairly deemed translatable into descriptions; witness 'John's mother'. Indeed, even so simple a term as 'Socrates', Russell has argued,[7] is for each of us synonymous with some description, perhaps '$(\imath x)$(x was a philosopher . x drank hemlock)' or perhaps another depending on how each of us first learned of Socrates. Are then all singular terms to be considered capable of fair translation into descriptions, except for those very few names which we may be supposed to have learned by direct confrontation with name and object? Must a separate category then be kept open for these few hypothetical exceptions?

Happily, we can isolate such epistemological considerations from the logic of singular terms by a very simple expedient: by insisting on the primacy of general terms. We may insist that what are learned by ostension, or direct confrontation, be never names but solely general terms. This we may insist on at the level strictly of logical grammar, without prejudice to epistemology or ontology. Without prejudice to epistemology because we may grant the epistemologist any of the words which he traces to ostension; we merely parse them differently. Instead of treating the ostensively learned word as a *name of* the shown object to begin with, we treat it to begin with as a general term *true* exclusively of the shown object; then we construe the name, as such, as amounting to '$(\imath x)Fx$'

[7] E.g., in "Knowledge by acquaintance."

where '*F*' represents that primitive general term. No matter to epistemology, but much to the smoothness of logical theory.

So there is no longer an obstacle to treating all singular terms as descriptions. Given any singular term of ordinary language, moreover, say 'Socrates' or 'Cerberus' or 'the author of *Waverley*', the proper choice of '*F*' for translation of the term into '$(\imath x)Fx$' need in practice never detain us. If a pat translation such as '$(\imath x)$(x wrote *Waverley*)' lies ready to hand, very well; if not, we need not hesitate to admit a version of the type of '$(\imath x)$(x is-Socrates)' or '$(\imath x)$(x is-Cerberus)', since any less lame version would, if admissible as a translation at all, differ at most in expository value and not in meaning.

Proofs of the type of the broker example set forth above are in no way facilitated by thus trivially transforming simple singular terms into descriptions. Construing 'John' in the broker example as a description '$(\imath x)$(x is-John)', or '$(\imath x)Jx$', would entitle us to a further descriptional premise '$\forall x(Jx \leftrightarrow .\ x = w)$', but this is neither necessary nor useful for the progress of the deduction. The advantage of treating all singular terms as descriptions is of a more theoretical kind: that of sparing us having to admit into the framework of our technical theory a distinction between a category of descriptions and a category of nondescriptive singular terms. It is theoretically important not to have to admit this distinction because, as we have seen, the question of there being essentially nondescriptive singular terms at all, and if so what, was shrouded in the theory of knowledge and meaning. We have segregated that issue from our concerns, by shifting it from the realm of singular terms into that of general terms. Every singular term can now, trivially if not otherwise, be handled as a description; what had been an issue over names learned ostensively versus names learned discursively now becomes an issue over general terms learned ostensively versus general terms learned discursively. In this form the issue ceases to cut across any of our schematism of logical forms and categories, and can be left to other minds.

There is a yet more striking benefit to be gained from treating all singular terms as descriptions, but it must await the next chapter.

EXERCISES

1. Express 'the tallest man in town' in the form '$(\imath x)(\ldots x \ldots)$', using 'taller than' but not 'tallest'.

2. Show that a descriptional premise, an axiom of identity, and the premise:

The author of *Waverley* wrote *Ivanhoe*

together imply:

Someone wrote both *Waverley* and *Ivanhoe*.

44
ELIMINATION OF SINGULAR TERMS

Let us next take up the problem, which has been looming for some time, of the truth value of such statements as 'Cerberus barks'. Falsity, as a sweeping answer covering all statements containing 'Cerberus', would be over-hasty: first, because the statement 'There is no such thing as Cerberus', at least, is true; and second, because whatever statements we adjudge false must admit of compounds, e.g., their negations, which will be true. Truth, as a sweeping answer, would encounter parallel difficulties.

Our deductive methods for singular terms throw no light on the question; for we already assume that the singular term names an object when we represent the singular term by a free variable, and we make the same assumption again when we adopt a descriptive premise for a description. Failing a named object, our methods show nothing, for what they purport to show rests then on a contrary-to-fact assumption. Common usage, moreover, likewise leaves us in the dark; for, excepting such contexts as 'There is no such thing as Cerberus', a singular term is ordinarily used only when the speaker believes or cares to pretend that the object exists.

Under ordinary usage, we saw (Chapter 3), truth values attach not to indicative conditionals as wholes but only to the consequents conditionally upon truth of the antecedents. Analogously, under ordinary usage truth values attach to contexts of singular terms for the most part only

conditionally upon existence of the objects. But if we are to have a smooth logical theory we must fill such gaps, even though arbitrarily, in such a way that every statement comes to have a truth value. Thus it was that we conventionally extended the concept of the conditional, in Chapter 3, so as to allow truth values generally to conditionals as wholes. An extension in the same spirit is needed now on the score of singular terms that do not name.

We cannot, we have seen, accomplish this extension by any blanket decision that all contexts of a term such as 'Cerberus' are to be false, or all true. We can, however, decide the simple contexts and then let the truth values of the compounds follow from those decisions. Let us speak of a general term as *simple,* for our purposes, when it does not explicitly have the form of an abstract or a negation, conjunction, alternation, conditional, or biconditional of shorter components. When any such simple general term is applied to a singular term which fails to name, let us classify the resulting sentence as false (for all values of any free variables it may have). Thus 'Cerberus barks', formed as it is by applying the simple term 'barks' to 'Cerberus', is adjudged false.

This rule is suited for use only upon sentences which are considered to be fully analyzed in point of logical structure. If the sentences are still to be subject to further paraphrasing of words into symbols, we must be wary of treating a predicate as "simple" in the sense above and then paraphrasing it into a complex one.

For illustration let us reexamine the statement:

(1) The broker who hired John hired only honors graduates.

If we use 'F' for 'broker' and 'G' for 'hired', then 'the broker who hired John' may be rendered '$(\imath x)(Fx \,.\, G\,x\,\text{John})$'. To say that this alleged person hired u is then to say:

$$(2) \qquad G\,(\imath x)(Fx \,.\, G\,x\,\text{John})\,u,$$

so that (1) becomes:

$$(3) \qquad \forall u[G\,(\imath x)(Fx \,.\, G\,x\,\text{John})\,u \rightarrow Hu]$$

where 'H' means 'honors graduate'. Now let us suppose that no broker or several hired John, so that there is no such thing as *the* broker who hired John. According to the decision which we have newly adopted to cover

such cases, the simple context (2) of '$(\imath x)(Fx \,.\, G\ x$ John)' is then to be classified as false for all choices of u. Thereupon the conditional in (3) becomes true for all choices of u through falsity of antecedent, so (3) becomes true. The outcome is therefore that (1) becomes true, independently of any consideration of honors graduates, in case John was hired by no broker or several. This particular outcome is the merest curiosity, neither welcome nor unwelcome, since ordinary usage leaves cases such as this undecided.

Even when a singular term fails to name, however, we do have very proper preconceptions about the truth value of the special context 'There is [or: is not] such a thing as . . .'. But statements of this form call for a separate analysis, along lines which are already pretty evident from these past observations:

(a) We may take '$(\imath x)Fx$' as the general form for singular terms.
(b) '$(\imath x)Fx$' purports to name the one and only object of which 'F' is true (supposing any general term for 'F' here).
(c) '$\forall x(Fx \leftrightarrow .\ x = y)$' amounts to saying that y is the one and only object of which 'F' is true.

To say that there is such a thing as $(\imath x)Fx$ is to say, in view of (b), that there is one and only one object of which 'F' is true; and this may, in view of (c), be said as follows:

(4) $$\exists y \forall x(Fx \leftrightarrow .\ x = y).$$

Here, then, is an adequate formulation of 'There is such a thing as $(\imath x)Fx$'; and no more can be wanted, in view of (a), in formulation of the general idiom 'There is such a thing as . . .'.

Curiously enough, the translation (4) of 'There is such a thing as $(\imath x)Fx$' is devoid of the singular term '$(\imath x)Fx$'. Now elimination of '$(\imath x)Fx$' from other contexts can also be accomplished. For, think of 'G' as representing any general term which is "simple" in the recently defined sense. Then '$G\ (\imath x)Fx$', which attributes 'G' to $(\imath x)Fx$, may be paraphrased as:

(5) $$\exists y[Gy \,.\, \forall x(Fx \leftrightarrow .\ x = y)].$$

This is seen as follows. First suppose (Case 1) that there is such a thing as $(\imath x)Fx$. Then the clause '$\forall x(Fx \leftrightarrow .\ x = y)$' identifies y with $(\imath x)Fx$, and

accordingly (5) as a whole becomes true or false according as 'G' is true or false of $(\imath x)Fx$. Next suppose (Case 2) that there is no such thing as $(\imath x)Fx$. Then '$\forall x(Fx \leftrightarrow .\ x = y)$' becomes false for all choices of y; so (5) becomes false. But '$G\ (\imath x)Fx$' likewise is to be false in this case, according to our recent agreement about simple general terms in application to singular terms that do not name.

We are now in a position to eliminate singular terms everywhere. Given any sentence involving singular terms, we begin by paraphrasing the sentence into the explicit notation of quantification and truth functions as fully as we can, leaving the singular terms undisturbed as components but putting each in the form of a description. Then we supplant each simple context of each description by its equivalent of the form (5)—or by (4) if it happens to have the form 'there is such a thing as $(\imath x)Fx$'.

For simplicity we have been imagining always *closed* singular terms, as opposed to open ones such as '$x + 5$' or 'the eldest son of x' or '$(\imath x)(x$ wrote $z)$'. Clearly, however, the open ones admit of elimination by the same procedure; the fact that a free variable is being carried along alters nothing essential to the reasoning.

To see how the elimination of singular terms proceeds in practice, let us return to (1) and eliminate the singular terms 'John' and 'the broker who hired John'. As a first step we may eliminate the description from the simple context (2). The general method of doing this was seen in the translation of '$G\ (\imath x)Fx$' into (5); but what we have to deal with now in place of '$G\ (\imath x)Fx$' is (2), which is of the form of '$G\ (\imath x)Fx$' with '$\{z\colon Fz\ .\ G\ z$ John$\}$' for 'F' and '$\{z\colon Gzu\}$' for 'G'. These same substitutions in (5) give:

(6) $$\exists y[Gyu\ .\ \forall x(Fx\ .\ G\ x \text{ John } .\leftrightarrow .\ x = y)],$$

then, as the translation of (2). But 'John' has yet to be dealt with. Writing 'J' for 'is-John', we render 'John' as a description '$(\imath z)Jz$', so that '$G\ x$ John' in (6) becomes '$Gx(\imath z)Jz$'. This clause is of the form '$G(\imath x)Fx$' with 'z' for 'x', 'J' for 'F', and '$\{z\colon Gxz\}$' for 'G'. Corresponding changes in (5) give:

(7) $$\exists w[Gxw\ .\ \forall z(Jz \leftrightarrow .\ z = w)],$$

then, as translation of '$Gx(\imath z)Jz$'. Now we have eliminated both singular terms. It remains only to assemble the pieces, by putting (7) for '$G\ x$ John' in (6) and then putting the result for (2) in (3). We thus get:

$$\forall u(\exists y[Gyu \,.\, \forall x(Fx \,.\, \exists w[Gxw \,.\, \forall z(Jz \leftrightarrow .\, z = w)] \\ .\leftrightarrow .\, x = y)] \rightarrow Hu)$$

as our final paraphrase of (1). But by now it begins to appear that the elimination of descriptions is of essentially theoretical interest, and that in practice the alternative handling of the broker problem which was noted in the preceding chapter recommends itself highly.

Nevertheless, the theoretical eliminability of singular terms—the dispensability of all names—is so startling that its importance scarcely needs dwelling upon except in the negative fashion of pointing out what it does not mean. It does not mean that our language loses all means of talking about objects; on the contrary, the foregoing considerations show that the extrusion of singular terms is unaccompanied by any diminution in the power of the language. What the disappearance of singular terms does mean is that all reference to objects of any kind, concrete or abstract, is channeled through general terms and bound variables. We can still say anything we like about any one object or all objects, but we say it always through the idioms of quantification: 'There is an object x such that . . .' and 'Every object x is such that . . .'. The objects whose existence is implied in our discourse are finally just the objects which must, for the truth of our assertions, be acknowledged as "values of variables"—i.e., be reckoned into the totality of objects over which our variables of quantification range. To be is to be a value of a variable. There are no ultimate philosophical problems concerning singular terms and their references, but only concerning variables and their values; and there are no ultimate philosophical problems concerning existence except insofar as existence is expressed by the quantifier '$\exists x$'. Except when we are concerned with philosophical issues of linguistic reference and existence, on the other hand, there is no point in depriving ourselves of the convenience of singular terms; and accordingly the techniques of inference hitherto developed for singular terms are not to be thought of as abandoned.

HISTORICAL NOTE: Frege had a notation for the description prefix in 1893. The present notation was adapted later from Peano by Russell. The important idea of eliminating descriptions by paraphrasing the context, substantially as above, was original with Russell, 1905; and it is in allusion to Russell's examples that the author of *Waverley* continues to be mentioned in subsequent writings on the subject. Russell did not take the further step of treating all names as descriptions and thus eliminating

them too. He preferred to preserve an epistemological distinction between names that were short for descriptions and names that were irreducibly proper, learned by acquaintance.

EXERCISE

Put the statement:

The woman who lives above us is German and loves flowers

into symbols, using 'Fx' for 'x is a woman', 'Gx' for 'x lives above us', 'Hx' for 'x is German', and 'Jx' for 'x loves flowers'. Then transform the whole so as to eliminate use of description.

45
ELIMINATION OF VARIABLES

We acquired in Chapters 18 and 19 an efficient Boolean logic of monadic terms. Terms were compounded to form further terms without help of abstraction, bound variables, or quantifiers. Terms were formed from terms by negation and conjunction and, dispensably, the further functors '$\vee$', '$\rightarrow$', and '$\leftrightarrow$'. Sentences were formed from terms by the existence functor '$\exists$' and, dispensably, the further functors '$\forall$', '$\subseteq$', '$\subset$', and '$\equiv$'.

Abstracts and bound variables were added afterward in preparation for quantification, which was wanted only in preparation for a logic of polyadic terms. It will now be shown that by generalizing negation, conjunction, and the existence functor to apply to all polyadic terms, and then adding four more functors, we can provide for a logic of polyadic terms that is just as powerful as the logic of abstracts or quantification but retains the spirit of our little Boolean logic of monadic terms. By this I mean that it has no bound variables, no abstracts, no quantifiers. The scheme will be explained by explaining how to translate back and forth

between it and the logic of variables and abstracts or, ultimately, of quantification. These translations serve explicitly to depict and analyze the combinatorial work that bound variables perform. It is this deepened understanding of the variable itself, rather than any practical advantages of a logic without variables, that makes the present excursion worth while.

Terms can be monadic or polyadic. An n-adic term will be said to be of *degree* n. In quantification theory, where the schematic term letters 'F', 'G', etc. appear only with variables attached, we can tell the degree of a term letter by counting the variables. Here, however, where variables are eventually to disappear, we shall still need to know the degrees of term letters; so let us adopt distinctive styles of term letters for different degrees. I shall use 'F^1', 'G^1', etc. as monadic letters, 'F^2', 'G^2', etc. as dyadic, and so on. Also for uniformity I shall use 'F^0', 'G^0', etc. instead of 'p', 'q', etc. as sentence letters; for sentences fit smoothly into the scheme as terms of degree 0.

This notation is not to be viewed as making logic newly dependent on number. The constant exponents are to be seen merely as parts of fancy letters. The numerical information that they convey is nothing more than was already available in quantification theory by counting the variables attached to a term letter.

Negation and conjunction, familiar thus far as applied to monadic terms and to sentences (hence terms of degree 0), are now to be generalized in obvious fashion to terms of all degrees. Resuscitating the polyadic term abstract (Chapter 28) for expository purposes, we may put the general case thus:

$$(1) \qquad \bar{F}^n \equiv \{x_1 x_2 \ldots x_n\colon -F^n x_1 x_2 \ldots x_n\},$$

$$(2) \qquad F^n G^n \equiv \{x_1 x_2 \ldots x_n\colon F^n x_1 x_2 \ldots x_n \cdot G^n x_1 x_2 \ldots x_n\}.$$

Incidentally the coextensiveness functor '$\equiv$' itself has undergone generalization at this point, joining not just monadic terms but n-adic ones for any n. Like the bound variables 'x_1', 'x_2', etc. and the term abstracts, however, '$\equiv$' enters here only as an expository device and not, like negation and conjunction, as part of the final notation.

The numerical variable 'n' in (1) and (2) will seem to have invoked numerical ideas unequivocally. So it has, but only, again, at the expository level and not within the notation that I am explaining. (1) and (2) are general explanations of '$\bar{F}^0$', 'F^0G^0', '$\bar{F}^1$', 'F^1G^1', and so on. These expressions are indeed schematic in turn; actual examples of '$\bar{F}^0$',

'F^1G^1', and '$\bar{F}^2$' would be '– (it is raining)', 'black dog', 'non-uncle of'. Numbers are not involved.

Negation and conjunction, previously at hand for terms of degree 0 (sentences) and 1, are generalized in (1) and (2) to apply to terms of any degree n. Now a generalization of '$\exists$' is likewise needed. Except for a momentary excursion in Chapter 28, where '$\exists$' was tentatively applied to dyadic abstracts, we have been applying it steadfastly to monadic terms and latterly only to monadic abstracts, rewriting the results as quantifications. The obvious generalization is the one suggested by that excursion in Chapter 28, whereby '$\exists$' when applied to an n-adic term would yield a term of degree 0 (a sentence). Consider, however, this different generalization: just as '$\exists$' when applied to a term of degree 1 yields a term of degree 0, so when applied to a term of degree n it yields a term of degree $n - 1$. This is the version that will serve us.

(3) $$\exists F^n \equiv \{x_2 \ldots x_n\colon \exists x_1 F^n x_1 x_2 \ldots x_n\}.$$

One of the four further functors wanted is *reflection,* the self-functor. Applied to the dyadic term 'hates', it yields the monadic term 'hates self'.

$$\mathrm{Ref}\, F^2 \equiv \{x\colon F^2 xx\}.$$

The generalization is this:

(4) $$\mathrm{Ref}\, F^n \equiv \{x_2 \ldots x_n\colon F^n x_2 x_2 \ldots x_n\}.$$

Another is *padding,* which goes poorly into words. It invokes gratuitous objects. Applied to 'dog', it produces a dyadic term 'Pad dog' that relates anything and everything to every dog.

$$\mathrm{Pad}\, F^1 \equiv \{xy\colon F^1 y\}.$$

The generalization is this:

(5) $$\mathrm{Pad}\, F^n \equiv \{x_0 x_1 \ldots x_n\colon F^n x_1 \ldots x_n\}.$$

The remaining two are functors of permutation, or inversion. The inevitable one in the dyadic case converts 'parent of' to 'offspring of' and active voice to passive. Its extension to n-adic terms could take either of

two natural lines. I shall need both, and will call them *major* and *minor inversion*.

(6) $\quad \text{Inv } F^n \equiv \{x_1 \ldots x_n\colon F^n x_n x_1 \ldots x_{n-1}\},$

(7) $\quad \text{inv } F^n \equiv \{x_1 \ldots x_n\colon F^n x_2 x_1 x_3 \ldots x_n\}.$

Iteration of 'Inv' serves to bring any variable to initial position. E.g.,

$$\begin{aligned} F^5vwxyz &\leftrightarrow (\text{Inv } F^5)wxyzv \\ &\leftrightarrow (\text{Inv Inv } F^5)xyzvw. \end{aligned}$$

But this maneuver disturbs the order of the rest of the variables; it dragged the old terminal variables '*yz*' forward. Now this disorder can be rectified by a second maneuver, which banishes a variable from second place to the far end. We apply 'inv' and then 'Inv', thus:

(8) $$\begin{aligned} G^5xyzvw &\leftrightarrow (\text{inv } G^5)yxzvw \\ &\leftrightarrow (\text{Inv inv } G^5)xzvwy. \end{aligned}$$

Repeating this maneuver, we can banish '*z*' in turn, ending up with the sought order '*xvwyz*'. Summed up,

$$F^5vwxyz \leftrightarrow (\text{Inv inv Inv inv Inv Inv } F^5)xvwyz.$$

What this shows is that repeated use of 'Inv' and 'inv' enables us to move any variable to initial position and still restore the order of the other variables. But this means that we can achieve any permutation whatever, building backward from right to left. Just pull the desired last variable to initial position, then pull the desired penultimate variable to initial position in front of it, and so on.

Invoking 'Ref', then, we can eliminate any repetition of variables. First we permute the repetition to initial position and then we apply 'Ref'; for

$$F^n x_2 x_2 \ldots x_n \leftrightarrow (\text{Ref } F^n)x_2 \ldots x_n.$$

By 'Pad' we can throw in any new variable, however irrelevant; for

$$F^n x_1 \ldots x_n \leftrightarrow (\text{Pad } F^n)x_0 x_1 \ldots x_n.$$

Afterward we can reposition the new variable as desired, by permutation.

Thus it is that our four added functors accomplish all the recombi-

natory work of variables. They suffice to *homogenize* any two predications—that is, to endow them with matching strings of variables, and indeed in any order, free of repetitions. For example, the heterogeneous predications 'F^5wzwxy' and 'G^4vxyz' are verifiably equivalent to these homogeneous ones:

(9) (Pad Ref Inv inv F^5)$vwxyz$, (inv Pad G^4)$vwxyz$.

Now I shall show that our seven functors, applied in iteration to term letters, are adequate to the whole logic of quantification. Bound variables go by the board. Functors and schematic term letters remain.

The translation proceeds as follows. Translate universal quantification into existential quantification and negation in the familiar way, and all truth functions into negation and conjunction. Put exponents on the term letters according to how many variables are attached, and change any sentence letters 'p', 'q', etc. to 'F^0', 'G^0', etc. Wherever a predication is negated, apply the negation rather to the term letter. Wherever two predications occur in conjunction, homogenize them as above and then reduce their conjunction to a predication of a conjunctive term; for, by (2),

$$F^n x_1 \ldots x_n \cdot G^n x_1 \ldots x_n \cdot \leftrightarrow (F^n G^n) x_1 \ldots x_n.$$

Continuing thus, we reduce each innermost quantification to an existential quantification of a single predication. Its term will be complex, usually, built up of term letters by our functors; and now we further complicate it by applying functors sufficient to bring the quantified variable into initial position without repetition. The quantification becomes '$\exists x \Gamma x y_1 \ldots y_n$', say, where '$\Gamma$' stands for some complex term; and this quantification reduces in turn, by (3), to '$(\exists \Gamma) y_1 \ldots y_n$'. The bound variable '$x$' is no more. As the translation continues, quantifications that were not innermost become so and are rid of their quantified variables in turn. All closed quantificational schemata, having no free variables, thus reduce to complex terms of degree 0 built up of term letters by our seven functors.

Further economy is available. George Myro showed me in 1971 that if instead of 'Inv' and 'inv' we adopt just a single permutation functor 'Perm' equivalent to the iteration 'Inv inv' illustrated in (8), then all the work of both 'Inv' and 'inv' can be done by 'Perm', '$\exists$', and 'Pad'.[8] Our seven functors thus reduce to six.

[8] This can be seen with help of *Ways of Paradox*, 1976, p. 298. The cropping functor there is our present '$\exists$'.

HISTORICAL NOTE: It was Schönfinkel in 1924 who first analyzed the combinatory work of variables by eliminating them in favor of combinatory functions. *Combinatory logic,* thus founded, has since accreted a substantial literature, primarily by H. B. Curry. That approach differs from the present one in that its combinators are not mere functors like ours, applicable to terms, but are terms in their own right, singular terms, names of functions. They are consequently applicable to themselves and one another, whereas the present functors apply only to terms to form new terms, and never apply to themselves or to one another. As a result the Schönfinkel-Curry notation is as powerful as that of higher set theory, and not readily shielded from the antinomies (*q.v. infra*), whereas the present notation is equivalent merely to that of quantification theory.

The present scheme dates from my "Variables explained away" (1960), apart from superficial details. It has some affinity with Tarski's "cylindrical algebras" (1952) and Halmos's "polyadic algebra" (1956), and more with a modified cylindrical algebra by Bernays (1959).

EXERCISES

1. Verify by explicit steps of transformation that 'F^5wzwxy' and 'G^4vxyz' are equivalent to (9).

2. Transform the schema '$\forall x(\exists y\, Fxy \rightarrow \forall y\, \exists z\, Gyxz)$' step by step into a complex term built of 'F^2' and 'G^3' by the term functors.

46
CLASSES

Confusion between a general term, true of each of various concrete objects, and a singular term naming an abstract object, is neither new nor uncommon. The same word is apt to play both roles. 'Red' is true of each red object and may also be said to name a property, a color. 'Man' is true of each man and may also be said to name a class, mankind.

Classes, attributes, and properties, if such there be, are abstract objects. Between attributes and properties I make no distinction, nor is there any call to distinguish classes from properties except on one point: classes are *extensional*. What this means is that they are counted identical when they have the same members. Properties, such as that of having a heart and that of having kidneys, may be properties of all the same individuals and yet be distinguished. What does make for identity of properties is not well defined; we are apt to be fobbed off with loose talk of essence versus accident, necessity versus contingency, or the like. So, despite the fact that 'property' or 'attribute' is more usual in everyday parlance than 'class', I favor the latter.

There are similarities in the behavior of general and singular terms that encourage confusion, and such confusion may well underlie the very belief in these abstract objects, whether classes or properties. Be that as it may, we must prize the outcome; for the assumption of properties, or better of classes, has implemented the theoretical resources of our language in ways indispensable to science and even to much unscientific communication. We shall presently note a few instances.

The confusion has shown itself in a readiness on the part of some logicians to let the schematic term letters '*F*', '*G*', etc. into quantifiers as if they were variables. What had been schematic letters for terms were thus bound as variables admitting objects of some sort as values—perhaps properties, better classes. At this point two conflations occur: general terms are conflated with class names, and schematic letters are conflated with variables.

I propose now to go along with the former of these conflations. It may have begun in confusion, but it can be followed without confusion. We can deliberately assign a term double duty, letting it be true of each man and at the same time name a class, mankind. We can let general terms double as names of their extensions. We shall find in Chapter 48 that we cannot assign this additional duty to every general term, because of antinomies; but we can assign it selectively.

Conflation of schematic letters with variables, on the other hand, profits us not at all and engenders only obscurity. If there are to be classes, let them be values of regular variables '*x*', '*y*', etc. like everything else. The schematic letter '*F*' continues to stand for general terms and, insofar as those terms are allowed now to double as class names, it stands as a schematic letter also for class names; but this still does not make it a bindable variable.

That there are men is written '$\exists x(\text{man } x)$'. That there *is* man—here

the term doubles as a class name—is written '$\exists x(x = \text{man})$'. That there are disembodied spirits is written:

$$\exists x(\text{disembodied } x \,.\, \text{spirit } x)$$

or, equivalently,

$$\exists \{x\colon \text{disembodied } x \,.\, \text{spirit } x\}$$

(see Chapter 22). That they constitute a class, on the other hand, is written:

$$\exists y(y = \{x\colon \text{disembodied } x \,.\, \text{spirit } x\}).$$

The set-theoretic notation of abstraction, appropriated in earlier chapters for the relative clause or term abstract, has now come home to roost. Like other general terms, it doubles now as a class name.

Epsilon was appropriated, along with the abstraction notation, as a copula alternative to juxtaposition for expressing predication: '$x \in F$' for 'Fx'. Little use was made of it then. Now that general terms are coming to do double duty as class names, epsilon likewise comes home to roost: it comes to express membership as well as predication. With classes as values of variables, indeed, it really comes to be needed; we can write '$x \in y$' where readers would be puzzled by the juxtaposition 'yx' though familiar with 'Fx'. When used thus before class names or variables, epsilon changes its grammatical status from that of a copula to that of a dyadic general term. In '$x \in y$' the general term epsilon is predicated of x and y.

What I have said of epsilon applies equally to the functors '$\subseteq$' and '$\subset$'. Traditionally they belonged to set theory, but when they emerged in Chapter 20 it was rather as copulas like epsilon. Now that general terms are coming to do double duty as class names, '$\subseteq$' and '$\subset$' come like epsilon to do double duty as dyadic general terms; thus '$x \subseteq y$', '$x \subset y$'.

The functor '$\equiv$' received in Chapter 20 the same status as '$\subseteq$' and '$\subset$': that of a copula for use between general terms. When general terms come to double as class names, '$\equiv$' takes on a second status correspondingly, coming like '$\subseteq$' and '$\subset$' to serve as a dyadic general term between variables or names of classes. But what dyadic general term? Precisely '$=$'! Accordingly I write '$x = y$', not '$x \equiv y$'; I write '$=$' rather than '$\equiv$' wherever we are unequivocally concerned with names rather than with general terms that may or may not double as class names.

Use of '$=$' rather than '$\equiv$' from the start, in Chapter 20, would have been more in keeping with my treatment of '$\in$', '$\subseteq$', and '$\subset$'. I could have gone on to explain that I was there using '$=$', like '$\in$', '$\subseteq$', and '$\subset$', as a copula rather than as a dyadic general term true of objects. This, however, would have been an excess of pedantic zeal and a hazard to clear thinking. It is safer to reserve '$=$' conspicuously to the identifying of objects and leave the equating of general terms to '$\equiv$'.

Up to a point the admitting of classes as values of variables can be explained away in terms merely of the consistency and validity of quantificational schemata. To affirm that

$$\forall y[\forall x(x \in y) \rightarrow \exists x(x \in y)],$$
$$\exists y[\exists x(x \in y) \,.\, -\forall x(x \in y)]$$

amounts merely to affirming the validity of the schema '$\forall x\, Fx \rightarrow \exists x\, Fx$' and the consistency of '$\exists x\, Fx \,.\, -\forall x\, Fx$'. Such an account is interesting in that, within its limits, it explains statements about classes without presupposing classes. Use is made merely of the concepts of quantificational validity and consistency. In Chapter 33 we found that these concepts could be specified in terms of deductive rules, or again in terms of substitution, without dependence on the class concept.

A statement about classes can be thus explained as long as its class quantifiers, as I shall call them, are all prenex and furthermore all its quantifiers, through its last class quantifier, are universal or all existential. (By a class quantifier here I mean merely any quantifier whose variable happens to recur after an epsilon.) If they are universal, truth of the statement amounts to validity of the corresponding quantificational schema; if they are existential, it amounts to consistency of the schema. E.g., truth of the statements:

$$\forall x \forall z[x \in z \,.\rightarrow\, \exists y(y \in z)], \qquad \exists z\, \exists w \forall x(x \in z \,.\leftrightarrow.\, x \in w)$$

of set theory amounts to validity and consistency of the respective schemata:

$$Fx \rightarrow \exists y\, Fy, \qquad \forall x(Fx \leftrightarrow Gx)$$

of quantification theory.

This expedient can be pushed a bit by bringing internal quantifiers into prenex position in the usual way. Thus the statement:

$$\exists z[\forall w \forall x(x \in w \mathrel{.\!\!\rightarrow\!\!.} x \in z) \rightarrow \exists y \exists x(x \in y \,.\, x \in z)],$$

with buried class quantifiers '$\forall w$' and '$\exists y$', can be transformed into:

$$\exists z \exists w \exists y[\forall x(x \in w \mathrel{.\!\!\rightarrow\!\!.} x \in z) \rightarrow \exists x(x \in y \,.\, x \in z)],$$

which can be explained as merely affirming in effect the consistency of the quantificational schema:

$$\forall x(Gx \rightarrow Fx) \rightarrow \exists x(Hx \,.\, Fx).$$

Laws of classes cannot be explained away in the above fashion, however, when their prenex quantifiers are mixedly universal and existential as in the examples:

$$\forall z \exists w \forall x(x \in z \mathrel{.\!\!\leftrightarrow\!\!.} x \in w),$$
$$\forall x \forall y \exists z(x \in z \mathrel{.\!\!\leftrightarrow\!\!.} y \in z).$$

It is in such statements that the irreducible substance of set theory, or the theory of classes, is to be sought. Similarly for statements containing buried class quantifiers which, when brought into prenex position by the rules of passage, become mixedly universal and existential; e.g.:

$$\forall x \forall y[\forall z(x \in z \mathrel{.\!\!\rightarrow\!\!.} y \in z) \rightarrow \forall w(y \in w \mathrel{.\!\!\rightarrow\!\!.} x \in w)].$$

But this power of expressing irreducibly new laws would of itself justify little interest in class theory, were it not accompanied by a corresponding increase of power on the side of application. A good example of this effect may be seen in the definition of the dyadic term 'ancestor' on the basis of 'parent'. To simplify the situation let us understand 'ancestor' in a slightly broadened sense, thereby counting as a person's ancestors not only his parents, grandparents, and so on, but also the person himself. Let us represent 'parent' by 'F', so that 'Fxy' means 'x is a parent of y'. Now the problem is to write 'x is an ancestor of y' using only 'F' and our various symbols.

An important feature of the class of y's ancestors is that all parents of members of the class are members of it in turn. Another feature of it is that y himself belongs to it. But these two features do not yet fix the class of y's ancestors uniquely; there are larger classes which also contain y and contain all parents of members. One such class is the class of the ancestors of y's grandsons. Another such class is the combined class of y's

ancestors and neckties; for, neckties being parentless, their inclusion does not disturb the fact that all parents of members are members. But clearly every class which contains y and all parents of members will have at least to contain all y's ancestors, no matter what extra things it may happen to contain. Moreover, *one* of these classes contains nothing but y's ancestors. Hence to be an ancestor of y it is necessary and sufficient to belong to every class which contains y and all parents of members. Therefore 'x is an ancestor of y' can be written thus:

> x belongs to every class which contains y and all parents of members

and schematized thus:

(1) $\forall u[y \in u \,.\, \forall z \forall w(w \in u \,.\, Fzw \,.\!\rightarrow.\, z \in u) \,.\!\rightarrow.\, x \in u]$.

This ingenious construction admits of many applications besides this genealogical one. An application to number will be seen in the next chapter. What is significant about the construction for present purposes is that it depends on quantifying over classes.

Another example of the power gained by quantifying over classes was proposed by Geach and perfected by David Kaplan, in private correspondence:

(2) Some people admire only one another.

Kaplan has proved that we cannot express this in terms of just 'person', 'admire', identity, truth functions, and quantification. (2) does not imply that there are two people who admire only each other. It might hold because of some irreducible group of say eleven mutual admirers. Invoking classes and membership, we can do justice to (2):

$$\exists z(\exists x(x \in z) \,.\, \forall x[x \in z \,.\!\rightarrow.\, \exists y(x \text{ admires } y) \,.\, \forall y(x \text{ admires } y \,.\!\rightarrow.\, x \neq y \,.\, y \in z)]).$$

A simpler illustration of our new access of power may be seen in the fact that the sign '=' of identity can now be defined; for things are identical just in case they belong to the same classes.

$$x = y \,.\!\leftrightarrow\, \forall z(x \in z \,.\!\rightarrow.\, y \in z).$$

We can then proceed to define also the universe, the empty class, the class whose sole member is x, and the class whose members are x and y.

$$\mathrm{V} = \{x: x = x\}, \qquad \Lambda = \{x: x \neq x\}.$$
$$\{x\} = \{y: y = x\}, \qquad \{x, y\} = \{z: z = x \mathbin{.\vee.} z = y\}.$$

HISTORICAL NOTE: The construction illustrated in the definition of ancestor was introduced by Frege in 1879 for application to number. It was rediscovered independently a few years later by Peirce, and again by Dedekind, who propounded it in 1887 under the name of the method of *chains*.

47
NUMBER

We say the Apostles are twelve, but not in the sense in which we say they are pious; for we attribute piety, but not twelveness, to each. 'The Apostles are pious' has the form '$\forall x(Fx \rightarrow Gx)$', with '$F$' for 'Apostle' and '$G$' for 'pious'; but 'The Apostles are twelve' has no such form, and is more nearly comparable to the mere existential quantification '$\exists x\, Fx$'. This familiar quantification may be read 'The Apostles are at least one'; and we might analogously think of 'The Apostles are twelve' as written '$\exists_{12} x\, Fx$', using a *numerically definite* quantifier.

These can be introduced on the basis purely of the theory of quantification and identity as of Chapter 42; there is no need here to assume classes. We can begin by explaining '$\exists_0 x\, Fx$':

$$\exists_0 x\, Fx \leftrightarrow\!\!- \exists x\, Fx.$$

Then we can explain each succeeding numerical quantifier in terms of its predecessor in a uniform way:

(1)
$$\exists_1 x\, Fx \leftrightarrow \exists x[Fx \,.\, \exists_0 y(Fy \,.\, y \neq x)],$$
$$\exists_2 x\, Fx \leftrightarrow \exists x[Fx \,.\, \exists_1 y(Fy \,.\, y \neq x)],$$

and so on. For anything to be true of $n + 1$ things is for it to be true of something other than which it is true of n things. By a dozen steps of expansion according to these definitions, '$\exists_{12} x\, Fx$' goes over into a schema of identity theory. Similarly for any other numerically definite quantifier. However, we still have no expansion of '$\exists_n x\, Fx$' with variable 'n'. We can say there are twelve Apostles and twelve Muses, in the form '$\exists_{12} x\, Fx \,.\, \exists_{12} x\, Gx$', but we find difficulty if we merely want to say that there are just as many Apostles as Muses without saying how many. No definitions are at hand for expanding:

$$\exists n(\exists_n x\, Fx \,.\, \exists_n x\, Gx).$$

All high-school students appreciate the persistence, and some the utility, of number variables in algebra. The example above and such related ones as 'There are twice as many eyes as faces' suggest that quantifiable number variables have a place also in the analysis of virtually unmathematical discourse.

If we are to have quantified variables for numbers we must find entities in our universe to view as numbers—or else expand our universe to include such entities. As a step toward a reasonable theory of numbers, consider the adjective 'twelvefold'. If we are to recognize such a term, and not just the corresponding quantifier '$\exists_{12} x$', we must recognize it as a term that is true not of persons, e.g. Apostles, but of classes, e.g. the class of Apostles. Rendered as a term abstract,

$$\text{twelvefold} \equiv \{y\colon \exists_{12} x(x \in y)\}.$$

The remaining step to the number 12 as an object is a short one: take 12 as the extension of 'twelvefold', hence as the class of all twelvefold classes. So, now that abstracts and other general terms are doubling as names of their extensions,

$$12 = \{y\colon \exists_{12} x(x \in y)\}.$$

Where z is a number, so construed, to say that a class y has z members is simply to say that $y \in z$.

The first of the numbers, 0, is then $\{\Lambda\}$: the class whose sole member is the empty class. 1 is then definable as the successor of 0, and 2 as the successor of 1, and so on, once we define successor. How to do that is evident from (1): the successor of z, call it Sz, is the class of all those classes which, when deprived of a member, become members of z. Thus

$$Sz = \{y\colon \exists x(x \in y \,.\, y - \{x\} \in z)\}.$$

Here $-\{x\}$ is the class of everything but x, so the conjunction $y - \{x\}$ is the class of all members but x of y.

Addition of numbers is easily defined in view of this circumstance: a class has $y + z$ members if and only if it can be broken into two parts having y and z members. But to say that a class w has z members is just to say that $w \in z$; so

$$y + z = \{x\colon \exists u\, \exists w(x = u \vee w \,.\, uw = \Lambda \,.\, u \in y \,.\, w \in z)\}.$$

Multiplication is definable in view of a similar circumstance: a class has $y \cdot z$ members if it can be broken into y parts having z members each. And only if? Not quite; there is an exception where $z = 0$, for we cannot have y empty classes if $y > 1$.[9] The formal definition of $y \cdot z$ is still a routine matter, but it runs long and we may pass over it.

What now of a formal definition of 'z is a number', 'Nz'? Numbers in the present sense are just 0, 1, 2, . . .—that is, 0 and the positive integers. Negative numbers, fractions, irrationals, and imaginaries are not of the sort used in measuring class size. So our present problem is to define 'Nz' in such a way that it will come out true when and only when z is 0 or 1 or 2, etc. A means of accomplishing this is suggested by the treatment of 'ancestor' in the preceding chapter. Just as 'ancestor of y' means 'y or parent of y or parent of parent of y or . . .', so 'number' means '0 or S0 or SS0 or . . .'. In close analogy to (1) of the preceding chapter, accordingly, we can define thus:

$$(2) \qquad Nz \leftrightarrow \forall u[0 \in u \,.\, \forall y(y \in u \,.\!\!\rightarrow\!. Sy \in u) \,.\!\!\rightarrow\!. z \in u].$$

To be a number is to belong to every class to which 0 belongs and the successor of each member belongs.

[9] Here I am indebted to Peter Geach.

This definition supports *mathematical induction,* a proof procedure central to number theory. We prove that $\forall z(\mathrm{N}z \rightarrow Fz)$ by proving that $F0$ and $\forall y(Fy \rightarrow F(\mathrm{S}y))$. The reader should observe how (2) justifies this inference.

We are now in a position to say that there are just as many Apostles as Muses.

$$\exists z(\mathrm{N}z \,.\, \text{Apostle} \in z \,.\, \text{Muse} \in z).^{10}$$

Note the use, here again, of general terms as names of their extensions.

Besides letting monadic general terms moonlight as names, and thus positing an abstract realm of classes, we might do the same for dyadic general terms. This would mean positing a second abstract realm, that of dyadic relations. At this point, however, we can economize; the classes already suffice for generating a full complement of relations or their reasonable facsimiles. A relation can be construed as a class of ordered pairs. The uncle relation is the class of all the uncle-niece and uncle-nephew pairs.

And what then is an ordered pair? Any arbitrary notion of a pair of objects x and y will serve our purposes perfectly so long as these conditions are met: (i) for any objects x and y there is such a pair; (ii) as soon as a pair is given, its first object x is thereby uniquely determined, and so is its second object y. (Thus the pair of x and y is different from the pair of y and x, unless $x = y$.) It happens that these conditions are met when the pair of x and y, written '$\langle x, y\rangle$', is defined arbitrarily thus:

$$(3) \qquad \langle x, y\rangle = \{\{x\}, \{x, y\}\}$$

—hence as the class of classes that has as members just the classes $\{x\}$ and $\{x, y\}$.

Dyadic term abstraction thereupon reduces to monadic, thus:

$$\{xy\colon \ldots x \ldots y \ldots\} \equiv \{z\colon \exists x \exists y(z = \langle x, y\rangle \,.\, \ldots x \ldots y \ldots)\}.$$

Like any monadic term abstract, it continues to function as a general term along with doing double duty as a name within such limits as may be imposed on class existence by the antinomies. In some cases

[10] Readers less gifted than yourself may feel at this point that the Muses are being treated more tolerantly than Cerberus (Chapter 41), thus failing to appreciate that our example is none the worse for being false.

$$\exists z(z = \{xy\colon \ldots x \ldots y \ldots\}),$$

in others not.

Like any monadic terms, dyadic ones are subject automatically to the Boolean functors. These may be thought of as defined by abstraction as in (8)–(12) of Chapter 21. Also the dyadic terms usher in a useful assortment of further functors. Obvious ones are those of inversion, or conversion, and of reflection; they simply repeat the 'inv F^2' and 'Ref F^2' of Chapter 45. Further there are the *Cartesian product,* the *resultant,* and the *image,* defined thus:

$$\begin{aligned} F \times G &\equiv \{xy\colon Fx \,.\, Gy\}, \\ F \,/\, G &\equiv \{xy\colon \exists z(Fxz \,.\, Gzy)\}, \\ F\text{`}\text{`}G &\equiv \{x\colon \exists y(Fxy \,.\, Gy)\}. \end{aligned}$$

Thus 'cat × dog' relates every cat to every dog; 'mother / parent' is equivalent to 'grandmother'; and the veteran`war are the veterans of wars.

A *correlation* is a relation that no one object bears to more than one, and no two objects bear to any one. Again we may as well adjust the definition to dyadic general terms without regard to whether they are names of relations.

$$\mathrm{Crln}\, F \leftrightarrow \forall x \forall y \forall z(Fxz \,.\, Fyz \,.\vee.\, Fzx \,.\, Fzy :\!\rightarrow.\, x = y).$$

This affords a second way of saying that there are just as many Apostles as Muses: they are in correlation. For saying that the F and the G are in correlation there is an old notation, 'sm' for 'similar':

$$F \;\mathrm{sm}\; G \;.\!\leftrightarrow\; \exists x(\mathrm{Crln}\, x \,.\, F \equiv x\text{`}\text{`}G \,.\, G \equiv (\mathrm{inv}\, x)\text{`}\text{`}F).$$

Numbers were treated above as sizes of classes. Being only 0 and the positive integers, however, they were the sizes only of finite classes. The *cardinal* numbers, so called in set theory, are more inclusive; they measure also infinite classes. Now just as the finite cardinal numbers were construed as classes of like-sized classes, so may the infinite ones. In general then the cardinal number of a class x, finite or infinite, is $\{y\colon y \;\mathrm{sm}\; x\}$. One might suppose that all infinite classes are alike in size and hence that there is at most one infinite number; and this would be to suppose that any two infinite classes are in correlation. Far from being

obvious, this can for generous universes be disproved; see *Set Theory and Its Logic*, §28.

Ways of construing numbers are available that are very unlike the one we have been studying. An elegant alternative, more widely followed nowadays, takes 0 simply as Λ and Sz as $z \vee \{z\}$. Each number thus becomes the class of all earlier ones. The definition (2) of 'Nz' then carries over, using the new versions of '0' and 'S'. This version of number also extends neatly into the infinite, but I shall not pursue it.

To say that y has z members was simply to say, in the earlier version of number, that $y \in z$. How would we say it now? The answer is evident when we reflect that each number z has z predecessors, counting 0, and hence, in the new version, z members. To say that y has z members is then to say that y sm z.

HISTORICAL NOTE: The definition of numerically definite quantification, in effect, is in Frege (1884). So is the interpretation of numbers as classes of appropriately sized classes, and so is the definition (2) of 'Nz'. The elegant alternative version of number noted in the last paragraphs is due to von Neumann (1923). The theory of infinite numbers is due to Cantor (1890). It was he who equated sizes of classes by correlation and proved, on certain reasonable assumptions in set theory, that there is no highest infinite.

The old algebra of relations stems from DeMorgan and Peirce and in part from Cayley's group theory (1854). See above, end of Chapter 27. Wiener in 1914 was the first to offer a generally applicable construction of the ordered pair. The version used above is a variant due to Kuratowski.

EXERCISES

1. Where

$$F \equiv \{xy\colon x \text{ is a man} \cdot y \text{ is a moment} \cdot x \text{ is alive in } y\},$$
$$G \equiv \{y\colon y \text{ is a moment in 1950}\}$$

put these four terms fairly naturally into words:

$$F\text{`}\text{`}G, \qquad F\text{`}\text{`}\bar{G}, \qquad -(F\text{`}\text{`}G), \qquad -(\bar{F}\text{`}\text{`}G).$$

2. Where 'F' and 'G' stand for 'admires' and 'brother of', put these seven terms fairly naturally into words:

$$-(F \mid G), \quad (\text{inv}\, F) \mid G, \quad F \mid (\text{inv}\, G), \quad (\text{inv}\, F) \mid (\text{inv}\, G),$$
$$\text{inv}((\text{inv}\, F) \mid G), \quad (\text{inv}\, \breve{F}) \mid G, \quad -((\text{inv}\, F) \mid (\text{inv}\, G)).$$

3. Find interpretations of 'F', 'G', 'H', and 'K' such that

$$-[(FG)\text{‘‘}H \equiv (F\text{‘‘}H)(G\text{‘‘}H)], \quad -[F\text{‘‘}(HK) \equiv (F\text{‘‘}H)(F\text{‘‘}K)].$$

Justify your answers.

4. Show that (3) fulfills the requirement (ii) on ordered pairs.

5. Spell out the justification by (2) of mathematical induction.

48
AXIOMATIC SET THEORY

We have been seeing samples of what can be expressed in set theory. The main sample comprised the finite cardinal numbers. One can go on to define the ratios, the irrational numbers, the imaginary numbers, functions, indeed the whole apparatus of classical mathematics, without exceeding the frugal notation of set theory. This notation, moreover, can be very frugal indeed. All that is needed besides quantification and the truth functions is the epsilon of membership. Abstraction is not needed. It can be defined by description thus:

$$(1) \qquad \{x\colon Fx\} = (\imath y)\forall x(x \in y \mathrel{.\leftrightarrow} Fx)$$

and description is eliminable as in Chapter 44. This definition covers abstraction only in its use as a class name and not in its broader use as a general term; but that broader use is heuristic and dispensable.

Schematic term letters turn up in our discourse *about* the notations of both logic and set theory, as just now, and terms of all sorts turn up in sentences that combine set theory with other topics in applying set theory

to the world. In actual sentences of set theory, however, there are no schematic letters—sentences never contain them—and furthermore there is no need of abstraction, thanks to (1), nor any call for any general term except epsilon itself. As noted in Chapter 46, epsilon figures in these contexts as a dyadic general term.

The reducibility of the whole vast conceptual apparatus of classical mathematics to this minuscule capsule is a reason for interest in set theory. Specialists are drawn also by an interest in the pure mathematics of the higher infinites, without regard to the translating or modeling of classical mathematics.

In these pages on set theory I have been occupied with questions of definition and not of theorems and proof, but I did mention that there are antinomies and that they stand in the way of treating all general terms as names of their extensions. The simplest and best known of these antinomies is *Russell's paradox*. By concretion ((13) of Chapter 21),

$$(1) \qquad y \in \{x\colon -(x \in x)\} \mathrel{.\leftrightarrow} -(y \in y),$$

but there is no such class; for, taking y as that class, we could infer from (1) that

$$y \in y \mathrel{.\leftrightarrow} -(y \in y)$$

—a contradiction. (1) is indeed true for every object y, but

$$(2) \qquad -\exists y(y = \{x\colon -(x \in x)\}).$$

Here, for all its negativity, is a firm little theorem of set theory.

The antinomy admits of variants and embellishments, kindred in varying degrees, and there is no way of neatly segregating them. There are pairs of abstracts such that either could consistently be taken to name a class, but not both. We cannot excise the lethal classes without excising innocuous ones as well, in one quarter or another. There are consequently many alternative axiom systems or proof procedures, differing from one another in respect of classes admitted and excluded. One system may excel in simplicity in some respects, another in naturalness in some respects. One system may be bolder and richer in its yield of theorems, another poorer and thus less likely to turn out inconsistent. The one in widest use by specialists today is Zermelo's or next of kin. I shall sketch Zermelo's.

Its most characteristic assumption of class existence is the axiom schema:

$$\exists z(z = \{y: y \in x \,.\, Fy\}),$$

which affirms the existence of any specifiable subclass of any given class x. Since it yields classes only from prior classes, further existence axioms are needed to prime the pump. We are given these:

$$\exists z(z = \{x, y\}),$$
$$\exists z(z = \{y: y \subseteq x\}),$$
$$\exists z(z = \{y: \exists w(y \in w \,.\, w \in x)\}).$$

A trend of the system is that a class exists unless it is too big. Consequently Zermelo cannot use the Fregean version of number studied in the preceding chapter; each number so construed, except 0, would be too big a class to exist. But the von Neumann version at the end of that chapter works very well.

Zermelo does need infinite classes—just not too excessively infinite. The foregoing axioms being too modest in that regard, he adds an axiom providing an infinite class. Further axioms trail off into the increasingly conjectural: the axiom of choice and the continuum hypothesis, to give names to two. There is no clear stopping place, since a complete system is not to be aspired to. That much is clear from Gödel's proof of the incompletability of number theory (Chapter 34), since number theory can be formulated in set theory.

Von Neumann contributed also a set theory of his own. It departed from Zermelo's by admitting some additional classes, large ones, in a marginal status: they were declared incapable of being members of further classes. Here it is that a use was found for the superfluity of terms, 'set' and 'class': 'set' is reserved for classes that can be members. That x is a set can be rendered '$\exists z(x \in z)$'. It seems that classes can be assumed without restriction as long as their members are restricted to sets, though von Neumann did not go that far. The burden of further existence axioms then comes in saying what classes are to qualify as sets. Von Neumann takes as sets Zermelo's classes. Classes are consequently sets for von Neumann if they are not too big.

In the course of these last three chapters we have inconspicuously but decisively taken leave of logic properly so called and moved into another department of mathematics, set theory. The transition occurs when the

question is answered, or even formulated, as to what classes there are: which ones among our general terms are to succeed in doubling as names. Because of the common failure to grasp the very distinction between general terms and class names, the crossing over from logic into set theory has often gone unheeded down the years. Much of set theory has been thought of vaguely as logic and conversely. Yet the distinction is sharp and significant. Set theory, like so much else in mathematics and like natural science as well, treats of its peculiar domain of objects. The logical truths, in contrast, are merely the sentences obtainable by substitution in valid logical schemata. They treat of no distinctive objects and indiscriminately of all.

HISTORICAL NOTE: The reduction of ratios to set theory is due to Peano and that of the irrationals is due chiefly to Dedekind (1872). The execution of the whole reduction program in detail came with Whitehead and Russell's monumental *Principia Mathematica* (1910–13), except for Wiener's step of reducing relations to classes. (See preceding historical note.)

Russell discovered his paradox in 1901. He propounded in 1908 his famous theory of types, a brand of set theory that commanded much favor for fifty years but would require excessive exposition for present purposes. Zermelo's system dates likewise from 1908, and von Neumann's from 1925. For fuller exposition and comparison of these systems and others, see my *Set Theory and Its Logic*.

PARTIAL ANSWERS TO EXERCISES

Chapter 1. *Ex. 2.*
The first three.

Ex. 4.
Yes, when 'p' and 'q' are both taken as false.

Chapter 2. *Ex. 1.*
The indicated routine gives:

$$-(pqr)\ -(\bar{p}\bar{q}r)\ -(\bar{p}q\bar{r})\ -(p\bar{q}\bar{r})\ -(\bar{p}\bar{q}\bar{r}).$$

But this is equivalent:

$$-(pqr)-(\bar{p}\bar{q})\ -(\bar{q}\bar{r})\ -(\bar{p}\bar{r}).$$

Chapter 3. *Ex. 2.*

$$-(p \vee q) \vee -(\bar{p} \vee \bar{q}).$$

Chapter 4. *Ex. 1.*
John will play or sing and Mary will sing.
John will play or John and Mary will sing.

Ex. 3, second part.

$$-(p \vee q) \cdot r \vee s \dashrightarrow -(p \vee q)s.$$

Ex. 5, last part.

$$\mathrm{EC}p\mathrm{KA}qr\mathrm{A}p\mathrm{K}qs\mathrm{CAKN}q\mathrm{N}r\mathrm{KN}p\mathrm{AN}q\mathrm{N}s\mathrm{N}p.$$

Chapter 5. *Ex. 2.*

$$\begin{array}{l} \bot\bot \vee \top \dashrightarrow \top\bot \\ \qquad \top \to \bot \\ \qquad\quad \bot \end{array}$$

Ex. 3, third part.

$$\begin{array}{l} \qquad\qquad p \to . p \to q \\ \top \to . \top \to q \quad \bot \to . \bot \to q \\ \quad \overline{\top \to q} \qquad\qquad \top \\ \qquad q \\ \quad \top \quad \bot \end{array}$$

Chapter 6. *Ex. 1, second part.*

$$p \leftrightarrow q \;.\mathsf{v}.\; p \leftrightarrow \bar{q}$$

$$\mathsf{T} \leftrightarrow q \;.\mathsf{v}.\; \mathsf{T} \leftrightarrow \bar{q} \qquad \bot \leftrightarrow q \;.\mathsf{v}.\; \bot \leftrightarrow \bar{q}$$

$$q \vee \bar{q} \qquad\qquad \bar{q} \vee q$$

$$\mathsf{T} \qquad\qquad \mathsf{T}$$

Ex. 2, first part.

$$p \vee q \;.\!\rightarrow q \;:\mathsf{v}:\; q \rightarrow .\; p \vee q.$$

Ex. 4.

(a) No. (b) Yes, the negation of '$p \vee q$' is consistent.
(c) Yes, the negation of '$p \vee \bar{p}$' is inconsistent.

Chapter 7. *Ex. 1.*

$$p \leftrightarrow .\, q \leftrightarrow r \;:\!\rightarrow:\; r \leftrightarrow .\, q \rightarrow p$$

$$\mathsf{T} \leftrightarrow .\, q \leftrightarrow r \;:\!\rightarrow:\; r \leftrightarrow .\, q \rightarrow \mathsf{T}$$

$$q \leftrightarrow r \;.\!\rightarrow r$$

$$q \leftrightarrow \mathsf{T} \;.\!\rightarrow \mathsf{T} \qquad q \leftrightarrow \bot \;.\!\rightarrow \bot$$

$$\mathsf{T} \qquad\qquad q$$

$$\mathsf{T} \quad \bot$$

So '$p \leftrightarrow .\, q \leftrightarrow r$' does not imply '$r \leftrightarrow .\, q \rightarrow p$'.

$$r \leftrightarrow .\, q \rightarrow p \;:\!\rightarrow:\; p \leftrightarrow .\, q \leftrightarrow r$$

$$r \leftrightarrow .\, q \rightarrow \mathsf{T} \;:\!\rightarrow:\; \mathsf{T} \leftrightarrow .\, q \leftrightarrow r$$

$$r \rightarrow .\, q \leftrightarrow r$$

$$\mathsf{T} \rightarrow .\, q \leftrightarrow \mathsf{T}$$

$$q$$

$$\mathsf{T} \quad \bot$$

So not vice versa.

Ex. 4.

Two fell swoops show that neither '$p \vee q \;.\!\rightarrow r$' nor '$p \rightarrow .\, q \vee r$' implies '$p \rightarrow q$'.

$$p \rightarrow q \qquad p \vee q \;.\!\rightarrow r \qquad p \rightarrow .\, q \vee r$$

$$\mathsf{T} \quad \bot \qquad \mathsf{T} \vee \bot \;.\!\rightarrow r \qquad \mathsf{T} \rightarrow .\, \bot \vee r$$

$$r \qquad\qquad r$$

Two fell swoops show that '$p \rightarrow q$' and '$p \vee q \;.\!\rightarrow r$' each imply '$p \rightarrow .\, q \vee r$'.

$$p \rightarrow .\, q \vee r \qquad p \rightarrow q \qquad p \vee q \;.\!\rightarrow r$$

$$\mathsf{T} \quad \bot \quad \bot \qquad \mathsf{T} \rightarrow \bot \qquad \mathsf{T} \vee \bot \;.\!\rightarrow \bot$$

$$\bot \qquad\qquad \bot$$

Truth-value analysis shows '$p \rightarrow q \;.\!\rightarrow:\; p \vee q \;.\!\rightarrow r$' not valid. So '$p \rightarrow q$' does not imply '$p \vee q \;.\!\rightarrow r$'. But it

would, by (ii), if '$p \rightarrow .\ q \vee r$' did, since '$p \rightarrow q$' implies '$p \rightarrow .\ q \vee r$'; so '$p \rightarrow .\ q \vee r$' does not imply '$p \vee q \ .\!\!\rightarrow r$' either.

Ex. 7, partial answer.
The schemata are '$p \leftrightarrow qr$' and '$q \rightarrow rp \ . \ \bar{q} \rightarrow \bar{r}\bar{p}$'.

Chapter 8. *Ex. 3.*
The 'then' clause extends through 'Barbuda', since the 'I'll' has to govern 'fly'. So the whole is a conditional:

> Either the Giants or the Bruins win and the Jackals take second place → I'll recover past losses and either buy a clavichord or fly to Barbuda.

The antecedent of this conditional is a conjunction and not an alternation, since 'the Giants' and 'the Bruins' are coordinated by the shared verb 'win'.

> The Giants or the Bruins win . the Jackals take second place .→ I'll recover past losses and either buy a clavichord or fly to Barbuda.

The consequent is likewise a conjunction, since its 'either' stands after the 'and' rather than before the 'I'll'. Final result:

> The Giants win ∨ the Bruins win . the Jackals take second place .→. I'll recover past losses . I'll buy a clavichord ∨ I'll fly to Barbuda.

Chapter 9. *Ex. 1.*
'$p \rightarrow r \ . \ q \rightarrow r$' is equivalent to '$p \vee q \ .\!\!\rightarrow r$', and the other three are equivalent to '$pq \rightarrow r$'. But the reader should produce the truth-value analyses.

Chapter 10. *Ex. 1, first schema.*
$\bar{p} \ . \ q \vee -[r \vee -(q \vee p)]$,
$\bar{p} : q \vee .\ \bar{r} \ . \ q \vee p$,
$\bar{p}q \vee \bar{p}\bar{r}q \vee \bar{p}\bar{r}p$.

This simplies to '$\bar{p}q \vee \bar{p}\bar{r}q$', and indeed to '$\bar{p}q$'. To see the advantage of working from the outside inward, try doing this exercise in the opposite way—first changing the innermost portion '$-(q \vee p)$' to '$\bar{q}\bar{p}$' and then working outward.

Ex. 1, last half.
The last three of the six schemata end up thus, after obvious simplifications:

$$pqr \vee \bar{p}\bar{q}\bar{r}, \qquad pqr \vee \bar{p}\bar{q}r \vee p\bar{q}\bar{r} \vee \bar{p}q\bar{r}, \qquad pqr \vee \bar{p}\bar{r} \vee \bar{q}\bar{r}.$$

But the reader should fill in all steps.

Ex. 2, last half.
Two of these last three have already gone into developed form, as seen above. The last one becomes:

$$pqr \vee \bar{p}q\bar{r} \vee \bar{p}\bar{q}\bar{r} \vee p\bar{q}\bar{r}$$

with omission of a repeated clause.

Ex. 3.
Yes. The reader should still justify it.

Ex. 4, last part (*corresponding to Ex. 4 of Chapter 7*).
The three schemata go into alternational normal form thus:

$$\bar{p} \vee q, \qquad \bar{p}\bar{q} \vee r, \qquad \bar{p} \vee q \vee r.$$

Developed, they eventually become respectively:

$$\bar{p}qr \vee \bar{p}\bar{q}r \vee \bar{p}q\bar{r} \vee \bar{p}\bar{q}\bar{r} \vee pqr \vee pq\bar{r},$$
$$\bar{p}\bar{q}r \vee \bar{p}\bar{q}\bar{r} \vee pqr \vee \bar{p}qr \vee p\bar{q}r,$$
$$\bar{p}qr \vee \bar{p}q\bar{r} \vee \bar{p}\bar{q}r \vee \bar{p}\bar{q}\bar{r} \vee pqr \vee pq\bar{r} \vee p\bar{q}r.$$

Rearranged for matching, they appear thus:

$$pqr \vee \bar{p}qr \vee \qquad \bar{p}\bar{q}r \vee p\bar{q}\bar{r} \vee \bar{p}q\bar{r} \vee \bar{p}\bar{q}\bar{r},$$
$$pqr \vee \bar{p}qr \vee p\bar{q}r \vee \bar{p}\bar{q}r \vee \qquad\qquad \bar{p}\bar{q}\bar{r},$$
$$pqr \vee \bar{p}qr \vee p\bar{q}r \vee \bar{p}\bar{q}r \vee p\bar{q}\bar{r} \vee \bar{p}q\bar{r} \vee \bar{p}\bar{q}\bar{r}.$$

Containment then shows that the third schema is implied by the first and by the second, and that there are no other implications. Clearly the method is poor.

Chapter 11. *Ex. 3, first part.*
We see by a fell swoop that 'qr' implies the rest of the alternation and so is redundant. Dropping it, we have '$p\bar{q}\bar{r} \vee \bar{p}q \vee pr$'. By a further fell swoop we see that this whole is implied by '$p\bar{q}$' and hence that the '$\bar{r}$' is still redundant.

Ex. 6, last half.
The last three schemata obtained in Ex. 1 of Chapter 10 yield no consensuses, since no two clauses are opposed in exactly one letter. So there are no further prime implicants.

Ex. 7, last half.
These three schemata have no shorter alternational normal equivalents, since all prime implicants are there and none is redundant.

Chapter 12. *Ex. 1.*
The first is dual to each of the other three, these three being equivalent. For, by the second law of duality, '$p \leftrightarrow q$' is dual to '$-(\bar{p} \leftrightarrow \bar{q})$'.

Ex. 3, clausal part.
The clause is redundant if a fell swoop shows it to be implied by the rest of the schema.

Ex. 4, first schema.

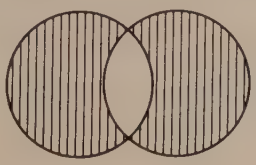
$$\bar{p} \cdot q \vee -[r \vee -(q \vee p)],$$
$$\bar{p} : q \vee . \bar{r} \cdot q \vee p,$$
$$\bar{p} \cdot q \vee \bar{r} \cdot q \vee q \vee p.$$

Simplifying,

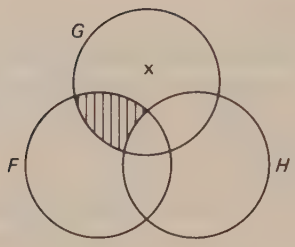
$$\bar{p} \cdot q \vee \bar{r} \cdot q \vee p$$
$$\bar{p} \cdot q \vee \bar{r} \cdot q,$$
$$\bar{p}q.$$

Chapter 13. *Ex. 1.*
By (1), $[3 \rightarrow :. 2 \rightarrow :] \bar{p} \rightarrow p \mathrel{.\!\rightarrow .} \bar{p} \rightarrow q$.

Chapter 15. *Ex. 2.*
Yes. There might be no F (and hence no G).



Chapter 17. *Ex. 1, second inference.*
Write 'F', 'G', and 'H' for 'knows George', 'knows Mabel', and 'admires Mabel'.



Ex. 3, second part.
Hint: Extend the bar device of Diagram 11

Chapter 18. *Ex. 1, first example.*
$-\exists FG\bar{H}, -\exists F\bar{H}\bar{G}, \therefore -\exists F\bar{H}.$

Ex. 1, last example.
$-\exists -(G \vee H), -\exists FG, \therefore \nvdash \exists F\bar{H}.$

Ex. 2.
It does not. (Reader should still explain, or, if baffled, read on.)

Chapter 19. *Ex. 2.*
'F' is 'takes logic', 'G' is 'takes Latin', and so on.

$$\exists FG\bar{H}\bar{J} \,.\, -\exists[(G \vee K) - (FJ)].$$

Negated:

$$-(\exists FG\bar{H}\bar{J} \,.\, -\exists[(G \vee K) - (FJ)]).$$

Converted to conjunctional normal form:

$$-\exists FG\bar{H}\bar{J} \vee \exists[(G \vee K) - (FJ)].$$

Converted to an existential conditional:

$$\exists FG\bar{H}\bar{J} \rightarrow \exists[(G \vee K) - (FJ)].$$

Fell swoop to test whether '$FG\bar{H}\bar{J}$' implies '$(G \vee K) - (FJ)$':

$$(G \vee K) - (FJ)$$
$$(\top \vee K) - (\top\bot)$$
$$\top$$

So the negation was valid. So the original was inconsistent.

Ex. 3, first part.

$$- \exists FG\bar{H} \,.\, - \exists F\bar{H}\bar{G} \,.\!\rightarrow\, - \exists F\bar{H}.$$

Conjunctional normal form:

$$\exists FG\bar{H} \vee \exists F\bar{H}\bar{G} \vee - \exists F\bar{H}.$$

Existential conditional:

$$\exists F\bar{H} \rightarrow \exists(FG\bar{H} \vee F\bar{H}\bar{G}).$$

Valid; for, fell swoop shows that '$F\bar{H}$' implies '$FG\bar{H} \vee F\bar{H}\bar{G}$'.

Chapter 20. *Ex. 1, second formula.*

$$\Lambda = (F \vee \bar{G})\bar{G} \vee -(F \vee \bar{G})G \,.\!\rightarrow\,. \Lambda = F\bar{G} \vee \bar{F}G,$$
$$- \exists[(F \vee \bar{G})\bar{G} \vee -(F \vee \bar{G})G] \rightarrow - \exists(F\bar{G} \vee \bar{F}G),$$
$$\exists(F\bar{G} \vee \bar{F}G) \rightarrow \exists[(F \vee \bar{G})\bar{G} \vee -(F \vee \bar{G})G].$$

Verify by truth-value analysis that '$F\bar{G} \vee \bar{F}G$' implies:

$$(F \vee \bar{G})\bar{G} \vee -(F \vee \bar{G})G.$$

Chapter 22. *Ex. 1, first example.*

$\forall x$(x is a man on the team $\rightarrow$ $-$ John can outrun x).

Ex. 2, first example.

$\forall x$(x is east of tracks $\rightarrow\!.$ x is slovenly $\vee$ x is poor),
$- \forall x$(x is east of tracks $\rightarrow$ x is poor),
$\therefore \exists x$(x is slovenly . $-$ x is poor).

Ex. 3, partial answer.
The two in the middle are not equivalent.

Chapter 23. *Ex. 2, partial answer.*
'$\forall x(p \leftrightarrow Fx)$' and '$p \leftrightarrow \forall x\, Fx$' resolve respectively to '$\forall x\ -Fx$' and '$-\forall x\, Fx$' when '$\perp$' is put for 'p', so they are not equivalent.

Ex. 3.
The first implies each and each implies the last. The reader should show this.

Chapter 24. *Ex. 1.*
Expanding '$\leftrightarrow$' and relettering, we have:

$$\forall x\, Fx \rightarrow\!.\, \forall y\, Gy \rightarrow \forall z\, Hz \,.\, \forall v\, Hv \rightarrow \forall w\, Gw.$$

Now we are free to choose the order of exportation of quantifiers. One of the final results is:

$$\exists x\,\exists y\,\exists v\forall z\forall w(Fx \rightarrow .\, Gy \rightarrow Hz\ .\ Hv \rightarrow Gw).$$

Note also the following refinement. We might proceed as far as the intermediate stage:

$$\exists x[Fx \rightarrow .\, \forall z(\forall y\ Gy \rightarrow Hz)\ .\ \forall w(\forall v\ Hv \rightarrow Gw)]$$

and then reletter thus:

$$\exists x[Fx \rightarrow .\, \forall z(\forall y\ Gy \rightarrow Hz)\ .\ \forall z(\forall v\ Hv \rightarrow Gz)]$$

with a view to exploiting the distributivity law (10) of Chapter 22, thus:

$$\exists x[Fx \rightarrow \forall z(\forall y\ Gy \rightarrow Hz\ .\ \forall v\ Hv \rightarrow Gz)].$$

The end result then has only four quantifiers:

$$\exists x\forall z\,\exists y\,\exists v(Fx \rightarrow .\, Gy \rightarrow Hz\ .\ Hv \rightarrow Gz).$$

Ex. 3.

$$\exists x\forall y(Fx \vee Gy \vee Hy\ .\ -Fx \vee -Gy \vee Hy),$$
$$\exists x[\forall y(Fx \vee Gy \vee Hy)\ .\ \forall y(-Fx \vee -Gy \vee Hy)],$$
$$\exists x[Fx \vee \forall y(Gy \vee Hy)\ .\ -Fx \vee \forall y(-Gy \vee Hy)],$$
$$\exists x[Fx\ .\ \forall y(-Gy \vee Hy)\ .\vee .\ -Fx\ .\ \forall y(Gy \vee Hy)$$
$$.\vee .\ \forall y(Gy \vee Hy)\ .\ \forall y(-Gy \vee Hy)],$$
$$\exists x[Fx\ .\ \forall y(-Gy \vee Hy)\ .\vee .\ -Fx\ .\ \forall y(Gy \vee Hy)]$$
$$\vee .\ \forall y(Gy \vee Hy)\ .\ \forall y(-Gy \vee Hy),$$
$$\exists x[Fx\ .\ \forall y(-Gy \vee Hy)] \vee \exists x[-Fx\ .\ \forall y(Gy \vee Hy)]$$
$$\vee .\ \forall y(Gy \vee Hy)\ .\ \forall y(-Gy \vee Hy),$$
$$\exists x\,Fx\ .\ \forall y(-Gy \vee Hy)\ .\vee .\ \exists x\ -Fx\ .\ \forall y(Gy \vee Hy)$$
$$.\vee .\ \forall y(Gy \vee Hy)\ .\ \forall y(-Gy \vee Hy).$$

Chapter 25. *Ex. 4.*

Hint: Represent the conclusion as:

$$\exists x(Fx\ .\ Gx) \rightarrow \exists x(Hx\ .\ Jx).$$

Chapter 26. *Ex. 2.*

$$\{z: y^z = zx + z\}.$$

Chapter 27. *Ex. 4.*

$$\forall x[x \text{ was a finger of hers}$$
$$\rightarrow \exists y(y \text{ was a ring}\ .\ y \text{ was on } x)].$$

Ex. 5, last example.
Hint: As a first step put just the inside quantification into words, retaining 'x'.

Ex. 6, last example.
False. If x and z were as described, then by taking y as z we could conclude that x was between z and z.

Ex. 7, first example.
'$- \exists x - \exists y(x = y)$', or, equivalently, '$\forall x - \forall y - (x = y)$', or, equivalently, '$\forall x \exists y(x = y)$'. True.

Chapter 28. *Ex. 1, left column.*
The first two may be substituted; the third not. The results of the first two substitutions mean that y praised himself to someone and that y praised someone to his face.

Ex. 2, partial answer.
Only the third abstract can be substituted in '$\forall x\, Fx \rightarrow Fy$' or in '$Fy \rightarrow \exists x\, Fx$'; only the third and fourth in (4); none in (13) or (14).

Ex. 3, partial answer.
Four of the eight are thus obtainable. The second on the left and the three on the right are not, although two of these could be inferred in an additional step or two.

Chapter 29. *Ex. 1, third part.*
Stages between (8) and (9):

$$\forall z(Fz \rightarrow Gz) \rightarrow \forall y[\exists x(Fx \,.\, Hyx) \rightarrow \exists w(Gw \,.\, Hyw)],$$
$$\forall y[\forall z(Fz \rightarrow Gz) \rightarrow .\, \exists x(Fx \,.\, Hyx) \rightarrow \exists w(Gw \,.\, Hyw)],$$
$$\forall y(\forall z(Fz \rightarrow Gz) \rightarrow .\, \forall x[Fx \,.\, Hyx \,.\!\rightarrow \exists w(Gw \,.\, Hyw)]),$$
$$\forall y \forall x[\forall z(Fz \rightarrow Gz) \rightarrow :\, Fx \,.\, Hyx \,.\!\rightarrow \exists w(Gw \,.\, Hyw)],$$
$$\forall y \forall x \exists z[Fz \rightarrow Gz \,.\!\rightarrow :\, Fx \,.\, Hyx \,.\!\rightarrow \exists w(Gw \,.\, Hyw)],$$
$$\forall y \forall x \exists z[Fz \rightarrow Gz \,.\!\rightarrow \exists w(Fx \,.\, Hyz \,.\!\rightarrow .\, Gw \,.\, Hyw)],$$
$$\forall y \forall x \exists z \exists w(Fz \rightarrow Gz \,.\!\rightarrow :\, Fx \,.\, Hyx \,.\!\rightarrow .\, Gw \,.\, Hyw).$$

Ex. 7.

$$\forall x \forall y \forall z(Fxy \,.\, Fyz \,.\!\rightarrow Fxz) \,.\, \forall x - Fxx \,.\!\rightarrow \forall x \forall y(Fxy \rightarrow - Fyx),$$
$$\forall x \forall y \forall z(Fxy \,.\, Fyz \,.\!\rightarrow Fxz) \,.\, \forall w - Fww \,.\!\rightarrow \forall u \forall v(Fuv \rightarrow - Fvu),$$

$\forall u \forall v \exists x \exists y \exists z \exists w (Fxy \,.\, Fyz .\!\rightarrow Fxz : -Fww$
$:\!\rightarrow .\, Fuv \rightarrow -Fvu),$
$\exists x \exists y \exists z \exists w (Fxy \,.\, Fyz .\!\rightarrow Fxz : -Fww$
$:\!\rightarrow .\, Fuv \rightarrow -Fvu),$
$Fuv \,.\, Fvu .\!\rightarrow Fuu : -Fuu :\!\rightarrow .\, Fuv \rightarrow -Fvu.$

Chapter 30. *Ex. 1.*
In 'Fyy' and '$-Fzy$' the instantial variable is not new. '$-\exists x\, Fxx$' is not prenex, but make it so and the proof goes through, with premises '$\forall x\, -Fxx$' and '$\forall y\, Fyy$' and instances as before.

Ex. 2, first part.
See beginning of Chapter 38.

Chapter 31. *Ex. 3, omitting some steps.*
$\forall x \forall y [Fx \,.\, Gxy .\!\rightarrow \exists z (Jzy \,.\, Hxz)],$
$\forall x (Fx \,.\, \exists y\, Gxy .\!\rightarrow \exists y\, Hxy).$

PRENEX PREMISES FOR INCONSISTENCY:
$\forall x \forall y \exists z (Fx \,.\, Gxy .\!\rightarrow .\, Jzy \,.\, Hxz)$
$\exists x\ \exists y\ \forall z\ -(Fx \,.\, Gxy .\!\rightarrow Hxz)$
INSTANCES:
$\forall z\ -(Fu \,.\, Guv .\!\rightarrow Huz)$ (double step)
$\exists z (Fu \,.\, Guv .\!\rightarrow .\, Jzv \,.\, Huz)$ (double step)
$Fu \,.\, Guv .\!\rightarrow .\, Jwv \,.\, Huw$
$-(Fu \,.\, Guv .\!\rightarrow Huw)$

Ex. 4.
For the results of paraphrasing, see answers to Chapter 35.

Ex. 5.
PREMISES:
$\forall x$(x laughs at $x \rightarrow$ I like x)
$\forall x[\forall y$(y is friend of $x \rightarrow x$ laughs at $y) \rightarrow$ I detest $x]$
PLATITUDE:
$\forall x$(I detest $x \rightarrow -$ I like x)
CONCLUSION:
$\exists x \forall y$(y is friend of $x \rightarrow x$ laughs at y)
$\rightarrow \exists x\ -\ x$ is friend of x

PRENEX PREMISES FOR INCONSISTENCY:
$\forall x(Gxx \rightarrow Lx)$
$\forall x \exists y(Fyx \rightarrow Gxy .\!\rightarrow Dx)$

$\forall x(Dx \rightarrow -Lx)$
$\exists x\ \forall y\ \forall z\ -(Fyx \rightarrow Gxy \mathrel{.\!\!\rightarrow} -Fzz)$

INSTANCES:

$\forall y\ \forall z\ -(Fyw \rightarrow Gwy \mathrel{.\!\!\rightarrow} -Fzz)$
$\exists y(Fyw \rightarrow Gwy \mathrel{.\!\!\rightarrow} Dw)$
$Fvw \rightarrow Gwv \mathrel{.\!\!\rightarrow} Dw$
$Gww \rightarrow Lw$
$Dw \rightarrow -Lw$
$-(Fvw \rightarrow Gwv \mathrel{.\!\!\rightarrow} -Fww)$ (double step)
$-(Fww \rightarrow Gww \mathrel{.\!\!\rightarrow} -Fww)$ (double step)

Chapter 32. *Ex. 2.*

The odd ones are T and the even ones $\perp$.

Chapter 34. *Ex. 1.*

There can be no complete disproof procedure here, since disproofs of negations afford proofs. There can be no decision procedure missing only finitely many of the truths, since that procedure together with a list of the omitted truths would afford a complete procedure.

Ex. 2.

Because otherwise you could always prove a schema to be finitely valid by proving that its negation was in B.

Chapter 35. *Ex. 1.*

Make substitutions to match those that were made in the first proof in this chapter, except for using 't_{uvu}' in (5).

Ex. 2.

Reworking Ex. 4 of Chapter 31 in Dreben's way:

PREMISES:

$\exists x(Fx\ .\ \forall y[Gy\ .\ \exists z(Fz\ .\ Hyz) \mathrel{.\!\!\rightarrow} Hyx])$
$\forall y[Gy \rightarrow \exists z(Fz\ .\ Hyz)]$
$-\exists x[Fx\ .\ \forall y(Gy \rightarrow Hyx)]$ (conclusion negated)

ADJUSTMENTS:

$\exists x(Fx\ .\ \forall y[-Gy \vee \forall z\ -(Fz\ .\ Hyz) \vee Hyx])$
$\forall y[-Gy \vee \exists z(Fz\ .\ Hyz)]$
$\forall x[-Fx \vee \exists y\ -(Gy \rightarrow Hyx)]$

FUNCTIONAL NORMAL FORMS:

$Fx\ .\ \forall y[-Gy \vee \forall z\ -(Fz\ .\ Hyz) \vee Hyx]$
$\forall y(-Gy \vee .\ Fz_y\ .\ Hyz_y)$
$\forall x[-Fx \vee -(Gy_x \rightarrow Hy_x x)]$

INSTANCES:

$Fx \,.\, -Gy_x \vee -(Fz_{y_x} \,.\, Hy_x z_{y_x}) \vee Hy_x x$
$-Gy_x \vee .\, Fz_{y_x} \,.\, Hy_x z_{y_x}$
$-Fx \vee -(Gy_x \rightarrow Hy_x x)$

Chapter 36. *Ex. 2.*
Proof of (4) of Chapter 29:

$Fxx \rightarrow Fxy \;.\vee.\; Fxy \rightarrow Fyy$
$\exists z(Fxz \rightarrow Fzy) \vee \exists z(Fxz \rightarrow Fzy)$
$\exists z(Fxz \rightarrow Fzy)$

Chapter 37. *Ex. 2.*
Proof of (1) of Chapter 29 by last method of Chapter 37:

[1] $Fy \,.\, \forall x(Fx \rightarrow Gxy) \rightarrow .\, Fy \,.\, Gyy$
$[\forall y\ 1 \rightarrow .]\ \exists y[Fy \,.\, \forall x(Fx \rightarrow Gxy)] \rightarrow$
$\exists x(Fx \,.\, Gxx)$

What valid monadic schema covers [1]?

$Fy \,.\, \forall x(Fx \rightarrow Hx) \;.\rightarrow.\; Fy \,.\, Hy.$

Proof of (4) of Chapter 29:

[1] $\forall z\, Fxz \rightarrow Fxy$
$[1 \rightarrow]\ \exists z(Fxz \rightarrow Fzy)$

The valid monadic schema that covers this last line is as follows. Test it for validity.

$\forall z\, Gz \rightarrow Hx \;.\rightarrow\; \exists z(Gz \rightarrow Hz).$

Proof of (8) of Chapter 29:

[1] $\forall x(Fx \rightarrow Gx) \rightarrow .\, \exists x(Fx \,.\, Hyx) \rightarrow \exists x(Gx \,.\, Hyx)$
$[\forall y\ 1 \rightarrow .]\ \forall x(Fx \rightarrow Gx) \rightarrow$
$\forall y[\exists x(Fx \,.\, Hyx) \rightarrow \exists x(Gx \,.\, Hyx)]$

Proof of (11) of Chapter 29:

[1] $\forall z(Fxz \rightarrow Fxz),$
$[1 \rightarrow]\ \exists y \forall z(Fxz \rightarrow Fyz).$

Chapter 38. *Ex. 1.*
The first and third on the left are correct but unfinished. The second on the left is incorrect (violating the alphabeti-

cal rule) and unfinished. The one at the upper right is correct and finished. The last one is finished but incorrect, 'y' being flagged twice.

Ex. 2 in part.
Revise the last one by bringing in a new line '$Fz \,.\, Gz$'.

Ex. 3.
Reworking Ex. 5 of Chapter 31 (whereof see answer above):

$*(1)$	$\forall x(Gxx \rightarrow Lx)$		
$**(2)$	$\forall x[\forall y(Fyx \rightarrow Gxy) \rightarrow Dx]$		
$***(3)$	$\forall x(Dx \rightarrow -Lx)$		
$****(4)$	$\exists x \forall y(Fyx \rightarrow Gxy)$		
$****(5)$	$\forall y(Fyw \rightarrow Gwy)$	(4)	w
$****(6)$	$Gww \rightarrow Lw$	(1)	
$****(7)$	$\forall y(Fyw \rightarrow Gwy) \rightarrow Dw$	(2)	
$****(8)$	$Dw \rightarrow -Lw$	(3)	
$****(9)$	$Fww \rightarrow Gww$	(5)	
$****(10)$	$-Fww$	(5)(6)(7)(8)(9)	
$****(11)$	$\exists x\ -Fxx$	(10)	
$***(12)$	$\exists x \forall y(Fyx \rightarrow Gxy) \rightarrow \exists x\ -Fxx$	$*(11)$	

Chapter 41. *Ex. 1 by method of pure existentials.*

$$-(Fy \,.\, Fz) \,.\, Fw \rightarrow \forall x\, Fx \,.\!\rightarrow -Fw,$$
$$\exists x[-(Fy \,.\, Fz) \,.\, Fw \rightarrow Fx \,.\!\rightarrow -Fw],$$
$$-(Fy \,.\, Fz) \,.\, Fw \rightarrow Fy \,.\!\rightarrow -Fw$$
$$:\!\vee\!: -(Fy \,.\, Fz) \,.\, Fw \rightarrow Fz \,.\!\rightarrow -Fw.$$

Ex. 2.

$$\forall x(Fxy \rightarrow Gyx) \,.\, \forall x(Gxz \rightarrow -Fzx) \,.\!\rightarrow -Fzy,$$
$$\exists x\, \exists w(Fxy \rightarrow Gyx \,.\, Gwz \rightarrow -Fzw \,.\!\rightarrow -Fzy),$$
$$Fzy \rightarrow Gyz \,.\, Gyz \rightarrow -Fzy \,.\!\rightarrow -Fzy.$$

This can be shortened by use of (10) of Chapter 22.

Chapter 42. *Ex. 1.*
Given that

$$\forall x(Fx \rightarrow\!: x = y \,.\vee.\, x = z), \qquad \exists x(Fx \,.\, Gx),$$

and given '$\forall x \forall w(Gx \,.\, x = w \,.\!\rightarrow Gw)$' as axiom of identity, to show that $Gy \vee Gz$. Main method:

PREMISES: $\forall x(Fx \rightarrow : x = y \mathbin{.\vee.} x = z)$
$\exists x(Fx \mathbin{.} Gx)$
$\forall x \forall w(Gx \mathbin{.} x = w \mathbin{.\rightarrow} Gw)$
$-(Gy \vee Gz)$

INSTANCES: $Fu \mathbin{.} Gu$
$Fu \rightarrow : u = y \mathbin{.\vee.} u = z$
$Gu \mathbin{.} u = y \mathbin{.\rightarrow} Gy$
$Gu \mathbin{.} u = z \mathbin{.\rightarrow} Gz$

The last five lines are truth-functionally inconsistent. The reader might try repeating the exercise by the method of pure existentials.

Ex. 2.
The forward conditional is a variant of (I), by a rule of passage. The reverse implication, when buttressed with the case '$y = y$' of (II), holds by EG.

Chapter 43. *Ex. 1.*

$(\imath x)[x$ is a man in town $\mathbin{.}$
$\forall y(y$ is a man in town $\mathbin{.} x \neq y \mathbin{.\rightarrow} x$ is taller than $y)]$.

Ex. 2 by method of pure existentials.

$\forall x(x$ wrote W $\leftrightarrow . x = y) \mathbin{.} y = y \mathbin{.} y$ wrote I
$\mathbin{.\rightarrow} \exists x(x$ wrote W $\mathbin{.} x$ wrote I),
$\exists x \exists z(x$ wrote W $\leftrightarrow . x = y : y = y \mathbin{.} y$ wrote I
$:\rightarrow . z$ wrote W $\mathbin{.} z$ wrote I),
y wrote W $\leftrightarrow . y = y : y = y \mathbin{.} y$ wrote I
$:\rightarrow . y$ wrote W $\mathbin{.} y$ wrote I.

Chapter 44.

$H(\imath x)(Fx \mathbin{.} Gx) \mathbin{.} J(\imath x)(Fx \mathbin{.} Gx)$,
$\exists y[Hy \mathbin{.} \forall x(Fx \mathbin{.} Gx \mathbin{.\leftrightarrow.} x = y)]$
$\mathbin{.} \exists y[Jy \mathbin{.} \forall x(Fx \mathbin{.} Gx \mathbin{.\leftrightarrow.} x = y)]$.

Chapter 47. *Ex. 1, last half.*
$-(F\text{`}\text{`}G)$ comprises everything, human and otherwise, except the men who were alive in 1950. $-(\breve{F}\text{`}\text{`}G)$ consists of the men who were born before 1950 and survived that year.

Ex. 2, fourth relation.
The relation of x to y where x is admired by someone (male or female) who has y as a brother.

Ex. 3, second half.
Take F as the benefaction relation, H as the class of Armenians, and K as the class of orphans. Then $F``(HK)$ comprises just the benefactors of Armenian orphans, whereas $(F``H)(F``K)$ includes also anyone else who benefits both an Armenian and an orphan.

BIBLIOGRAPHY

This list includes only such logical and nearly logical works as happen to have been alluded to, by title or otherwise, in the course of the book. For a comprehensive register of the literature of mathematical logic to the end of 1935 see Church's *Bibliography*. That work, which is helpfully annotated and thoroughly indexed by subjects, is invaluable to logicians. Subsequent literature is covered by the Reviews section of the *Journal of Symbolic Logic,* which is indexed by subjects every five years and by authors every two.

AIKEN, Howard H., and others. *Synthesis of Electronic Computing and Control Circuits*. Cambridge, Mass.: Harvard University Press, 1951.

BEHMANN, Heinrich. *Mathematik und Logik*. Leipzig, 1927.

——— "Beiträge zur Algebra der Logik, insbesondere zum Entscheidungsproblem." *Mathematische Annalen,* vol. 86 (1922), pp. 163–229.

BERNAYS, Paul. "Über eine natürliche Erweiterung des Relationenkalküls," in A. Heyting, editor, *Constructivity in Mathematics* (Amsterdam: North-Holland, 1959), pp. 1–14.

——— and Moses Schönfinkel. "Zum Entscheidungsproblem der mathematischen Logik." *Mathematische Annalen,* vol. 99 (1928), pp. 342–372.

BOOLE, George. *The Mathematical Analysis of Logic*. London: George Bell, and Cambridge, Eng.: Macmillan, 1847. Reprinted in *Collected Logical Works,* Oxford: Blackwell, and New York: Philosophical Library, 1948.

——— *An Investigation of the Laws of Thought*. London: Walton and Maberly, 1854. Reprinted New York: Dover, 1951.

CANTOR, Georg. "Über eine elementare Frage der Mannigfältigkeitslehre." *Jahresberichte* der deutschen Mathematiker-Vereinigungen, vol. 1 (1890–91), pp. 75–78. Reprinted in *Gesammelte Abhandlungen,* Berlin: Springer, 1932.

CARNAP, Rudolf. *The Logical Syntax of Language*. London and New York: Harcourt Brace, 1937.

CARROLL, Lewis. *Symbolic Logic*. London: Macmillan, 1897.

CAYLEY, Arthur. "On the theory of groups as depending on the symbolical equation $\theta^n = 1$." *Philosophical Magazine,* vol. 7 (1854), pp. 40–47, 408–409. Reprinted in *Collected Mathematical Papers,* Cambridge, Eng.: Cambridge University Press, 1895.

CHURCH, Alonzo. *A Bibliography of Symbolic Logic.* Providence, R.I.: Association for Symbolic Logic, 1938. Reprinted from *Journal of Symbolic Logic* (1936, 1938).

——— "A note on the Entscheidungsproblem." *Journal of Symbolic Logic,* vol. 1 (1936), pp. 40–41. Correction on pp. 101–102.

COOLEY, John C. *A Primer of Formal Logic.* New York: Macmillan, 1942.

CURRY, H. B., and Robert Feys. *Combinatory Logic.* Amsterdam: North-Holland, 1958.

DAVIS, Martin. *Computability and Unsolvability.* New York: McGraw-Hill, 1958.

———, ed. *The Undecidable.* Hewlett, N.Y.: Raven Press, 1965.

DEDEKIND, Richard. *Stetigkeit und irrationale Zahlen.* Braunschweig: Vieweg, 1872. Later editions 1892, 1905, 1912.

——— *Was sind und was sollen die Zahlen?* Braunschweig: Vieweg, 1887.

DEMORGAN, Augustus. *Formal Logic.* London: Taylor, 1847.

——— "On the syllogism." *Transactions* of the Cambridge Philosophical Society, vol. 8 (1849), pp. 379–408; vol. 9 (1856), pp. 79–127; vol. 10 (1864), pp. 173–230, 331–358, 428–487. These had been read at earlier meetings: 1846, 1850, 1860, 1863.

DREBEN, Burton. "On the completeness of quantification theory." *Proceedings* of the National Academy of Sciences, vol. 38 (1952), pp. 1047–52.

FEIGL, Herbert, and Wilfrid Sellars, eds. *Readings in Philosophical Analysis.* New York: Appleton-Century-Crofts, 1949.

FREGE, Gottlob. *Begriffsschrift.* Halle: Nebert, 1879. Translated in van Heijenoort.

——— *Die Grundlagen der Arithmetik.* Breslau: Koebner, 1884. Reprinted with English translation, Oxford: Blackwell, and New York: Philosophical Library, 1950.

——— *Grundgesetze der Arithmetik.* Vol. 1, 1893; vol. 2, 1903. Jena: Pohle.

——— "Über Sinn und Bedeutung." *Zeitschrift für Philosophie und philosophische Kritik,* vol. 100 (1892), pp. 25–50. Translated in Feigl and Sellars.

FRIDSHAL, R. Abstract in *Summaries of Talks at the Summer Institute of Symbolic Logic*. Mimeographed. Ithaca, N.Y.: Cornell University, 1957, pp. 211–212.

GENTZEN, Gerhard. "Untersuchungen über das logische Schliessen." *Mathematische Zeitschrift,* vol. 39 (1934–35), pp. 176–210, 405–431.

GERGONNE, J. D. "Essai de dialectique rationelle." *Annales de mathématiques pures et appliquées,* vol. 7 (1816–17), pp. 189–228.

GHAZALA, M. J. "Irredundant disjunctive and conjunctive forms of a Boolean function." *I.B.M. Journal of Research and Development,* vol. 1 (1957), pp. 171–176.

GÖDEL, Kurt. *The Consistency of the Continuum Hypothesis*. Princeton, N.J.: Princeton University Press, 1940.

——— "Die Vollständigkeit der Axiome des logischen Funktionenkalküls." *Monatshefte für Mathematik und Physik,* vol. 37 (1930), pp. 349–360. Translated in van Heijenoort.

——— "Über formal unentscheidbare Sätze der Principia Mathematica und verwandter Systeme." *Monatshefte für Mathematik und Physik,* vol. 38 (1931), pp. 173–198. Translated in van Heijenoort.

——— "Zum Entscheidungsproblem des logischen Funktionenkalküls." *Monatshefte für Mathematik und Physik,* vol. 40 (1933), pp. 433–443.

——— "On undecidable propositions of formal mathematical systems." Mimeographed. Princeton, N.J., 1934. Printed in Davis, *The Undecidable*.

GOODMAN, Nelson. "The problem of counterfactual conditionals." *Journal of Philosophy,* vol. 44 (1947), pp. 113–128. Reprinted in *Fact, Fiction, and Forecast,* Indianapolis, Ind.: Bobbs-Merrill, 1965.

HALMOS, Paul. "Algebraic logic (II)." *Fundamenta Mathematicae,* vol. 43 (1956), pp. 255–325.

HERBRAND, Jacques. *Recherches sur la théorie de la démonstration*. Warsaw: Société des Sciences et des Lettres, 1930. Translated in Herbrand, *Logical Writings,* Cambridge, Mass.: Harvard University Press, 1971.

HILBERT, David, and W. Ackermann. *Grundzüge der theoretischen Logik*. Berlin: Springer, 1928. 2d ed., 1938. 3d ed., 1949.

——— and P. Bernays. *Grundlagen der Mathematik*. Vol. 1, 1934; vol. 2, 1939. Berlin: Springer.

JAŚKOWSKI, Stanisław. "On the rules of suppositions in formal logic." *Studia Logica,* no. 1 (Warsaw, 1934).

JEVONS, W. Stanley. *Pure Logic*. London: Stanford, 1864.

KLEENE, S. C. *Introduction to Metamathematics*. Amsterdam and New York: Van Nostrand Reinhold, 1952.

——— "Recursive predicates and quantifiers." *Transactions* of the American Mathematical Society, vol. 53 (1943), pp. 41–73.

KURATOWSKI, Casimir. "Sur la notion de l'ordre dans la théorie des ensembles." *Fundamenta Mathematicae,* vol. 2 (1921), pp. 161–171.

LEŚNIEWSKI, Stanisław. "Grundzüge eines neuen Systems der Grundlagen der Mathematik." *Fundamenta Mathematicae,* vol. 14 (1929), pp. 1–81.

LEWIS, C. I. *A Survey of Symbolic Logic*. Berkeley: University of California Press, 1918.

LÖWENHEIM, Leopold. "Über Möglichkeiten im Relativkalkül." *Mathematische Annalen,* vol. 76 (1915), pp. 447–470. Translation in van Heijenoort.

ŁUKASIEWICZ, Jan. "O logice trójwartościowej" (On three-valued logic). *Ruch Filozoficzny,* vol. 5 (1920), pp. 169–171.

——— "Uwagi o aksyomacie Nicod'a i o dedukcyi uogólniajacej" (Remarks on Nicod's axiom and on generalizing by deduction). *Ksiega Pamiatkowa Polskiego Towarzystwa Filozoficznego we Lwowie,* Lwów, 1931.

——— "Zur Geschichte der Aussagenlogik." *Erkenntnis,* vol. 5 (1935–36), pp. 111–131.

MARTIN, R. M. "A homogeneous system for formal logic." *Journal of Symbolic Logic,* vol. 8 (1943), pp. 1–23.

NEUMANN, J. von. "Zur Einführung der transfiniten Zahlen." *Acta Litterarum ac Scientiarum Regiae Universitatis Hungaricae Francisco-Josephinae* (sect. scient. math.), vol. 1 (Szeged, 1923), pp. 199–208. Translation in van Heijenoort.

——— "Eine Axiomatisierung der Mengenlehre." *Journal für reine und angewandte Mathematik,* vol. 154 (1925), pp. 219–240; vol. 155 (1926), p. 128. Translation in van Heijenoort.

NICOD, Jean. "A reduction in the number of primitive propositions of logic." *Proceedings* of the Cambridge Philosophical Society, vol. 19 (1917–20), pp. 32–44.

PEANO, Giuseppe. *Arithmetices Principia*. Turin, 1889.

——— *Formulaire de Mathématiques*. Introduction, 1894; vol. 1, 1895; vol. 2, 1897–99. Turin. Vol. 3, 1901, Paris. Vol. 4, 1902–3; vol. 5 (s.v. *Formulario Matematico*), 1905–8. Turin: Bocca.

PEIRCE, C. S. *Collected Papers,* vols. 2–4. Cambridge, Mass.: Harvard University Press, 1932–33.

POST, E. L. "Introduction to a general theory of elementary propositions." *American Journal of Mathematics,* vol. 43 (1921), pp. 163–185. Reprinted in van Heijenoort.

PRESBURGER, M. "Über die Vollständigkeit eines gewissen Systems der Arithmetik." *Sprawozdanie* z I Kongresu Matematyków Krajów Slowánskych (Warsaw, 1930), pp. 92–101, 395.

QUINE, W. V. *Mathematical Logic.* New York, 1940. 2d printing, Cambridge, Mass.: Harvard University Press, 1947. Rev. ed., 1951.

——— *Elementary Logic.* Boston, 1941. Rev. ed., Cambridge, Mass.: Harvard University Press, 1966.

——— *O Sentido da nova lógica.* São Paulo: Martins, 1944.

——— *From a Logical Point of View.* Cambridge, Mass.: Harvard University Press, 1953.

——— *Set Theory and Its Logic.* Cambridge, Mass.: Harvard University Press, 1963, 1969.

——— *Selected Logic Papers.* New York: Random House, 1966.

——— *The Ways of Paradox and Other Essays,* enlarged edition. Cambridge, Mass.: Harvard University Press, 1976.

——— "Variables explained away." *Proceedings* of the American Philosophical Society, vol. 104 (1960), pp. 343–347. Reprinted in *Selected Logic Papers.*

ROGERS, Hartley, Jr. *Theory of Recursive Functions and Effective Computability.* New York: McGraw-Hill, 1967.

RUSSELL, Bertrand. *The Principles of Mathematics.* Cambridge, Eng.: Cambridge University Press, 1903. 2d ed., New York, 1938.

——— "On denoting." *Mind,* vol. 14 (1905), pp. 479–493. Reprinted in Feigl and Sellars.

——— "Mathematical logic as based on the theory of types." *American Journal of Mathematics,* vol. 30 (1908), pp. 222–262.

——— "Knowledge by acquaintance and knowledge by description." *Proceedings* of the Aristotelian Society, vol. 11 (1911), pp. 108–128. Reprinted in *The Problems of Philosophy,* London: Williams and Norgate, and New York: Holt, 1912, and in *Mysticism and Logic,* New York: Longmans, 1918.

SAMSON, E. W., and B. E. Mills. "Circuit minimization: algebra and algorithms for new Boolean canonical expressions." *ARCRC Technical Report* 21, 1954.

SCHÖNFINKEL, Moses. "Ueber die Bausteine der mathematischen Logik." *Mathematische Annalen,* vol. 92 (1924), pp. 305–316. Translation in van Heijenoort.

SCHRÖDER, Ernst. *Der Operationskreis des Logikkalküls*. Leipzig: Teubner, 1877 (37 pp.).

——— *Vorlesungen über die Algebra der Logik*. Vol. 1, 1890; vol. 2, 1891–1905; vol. 3, 1895. Leipzig: Teubner.

SHANNON, Claude E. "A symbolic analysis of relay and switching circuits." *Transactions* of the American Institute of Electrical Engineers, vol. 57 (1938), pp. 713–723.

SHEFFER, H. M. "A set of five independent postulates for Boolean algebras." *Transactions* of the American Mathematical Society, vol. 14 (1913), pp. 481–488.

SHOENFIELD, Joseph R. *Mathematical Logic*. Reading, Mass.: Addison-Wesley, 1967.

SKOLEM, Thoralf. "Logisch-kombinatorische Untersuchungen über die Erfüllbarkeit oder Beweisbarkeit mathematischer Sätze." *Skrifter* utgitt av Det Norske Videnskaps-Akademi i Oslo, I. Mat.-naturv. Kl. 1920, no. 4. Translation in van Heijenoort.

——— "Über die mathematische Logik." *Norsk Matematisk Tidsskrift*, vol. 10 (1928), pp. 125–142. Translation in van Heijenoort.

——— "Über einige Satzfunktionen in der Arithmetik." *Skrifter* utgitt av Det Norske Videnskaps-Akademi i Oslo, I. Mat.-naturv. Kl. 1930, no. 7.

SMULLYAN, Raymond M. *Theory of Formal Systems*. Princeton, N.J.: Princeton University Press, 1961.

TARSKI, Alfred. *Introduction to Logic*. New York: Oxford University Press, 1941.

——— *A Decision Method for Elementary Algebra and Geometry*. Santa Monica: Rand, 1948. Rev. ed., Berkeley: University of California Press, 1951.

——— *Logic, Semantics, Metamathematics*. Oxford: Clarendon Press, 1956.

——— "Remarques sur les notions fondamentales de la méthodologie des mathématiques." *Rocznik* Polskiego Towarzystwa Matematycznego, vol. 7 (1929), pp. 270–272.

———, A. Mostowski, and R. M. Robinson. *Undecidable Theories*. Amsterdam: North-Holland, 1953.

——— and F. B. Thompson. Abstracts in *Bulletin* of American Mathematical Society, vol. 58 (1952), pp. 65f.

TRACHTENBROT, B. A. "Impossibility of an algorithm for the decision problem in finite classes" (Russian). *Doklady* Akademii Nauk SSSR, vol. 70 (1950), pp. 569–572. Reviewed in *Journal of Symbolic Logic*, vol. 15 (1950), p. 229.

TURING, Alan M. "On computable numbers, with an application to the Entscheidungsproblem." *Journal* of the London Mathematical Society, vol. 42 (1936–37), pp. 230–265. Correction in vol. 43, pp. 544–546.

VAN HEIJENOORT, Jean, ed. *From Frege to Gödel: A Source Book in Mathematical Logic*. Cambridge, Mass.: Harvard University Press, 1967.

VENN, John. *Symbolic Logic*. London: Macmillan, 1881. 2d ed., 1894.

——— "On the diagrammatic and mechanical representations of propositions and reasoning." *The London Edinburgh, and Dublin Philosophical Magazine and Journal of Science,* vol. 10 (1880), pp. 1–18.

VON NEUMANN, see Neumann.

WHITEHEAD, A. N., and B. Russell. *Principia Mathematica,* vol. 1, 1910; vol. 2, 1912; vol. 3, 1913. Cambridge, Eng.: Cambridge University Press, 2d ed., 1925–27.

WIENER, Norbert. "A simplification of the logic of relations." *Proceedings* of the Cambridge Philosophical Society, vol. 17 (1912–14), pp. 387–390. Reprinted in van Heijenoort.

WITTGENSTEIN, Ludwig. *Tractatus Logico-Philosophicus*. London: Kegan Paul, 1922. Reprint of "Logisch-philosophische Abhandlung" (*Annalen der Naturphilosophie,* 1921) with English translation in parallel.

ZERMELO, Ernst. "Untersuchungen über die Grundlagen der Mengenlehre." *Mathematische Annalen,* vol. 65 (1908), pp. 261–281. Translation in van Heijenoort.

INDEX

A, *see* Categorical
∀, 128–130, 137, 141
Absolute term, 137, 167
Abstract object, 5, 94, 259–261, 289. *See also* Class
Abstraction, 134–136, 161; polyadic, 175–176, 284, 297–298; as description, 300
Ackermann, W., 120, 181
Aiken, H. H., 78
Algebra: of subtraction, 87–88; of classes, 114, 130–132; of relations, 173; of real numbers, 218, 295
Alphabetical order, 127
Alphabetical restriction, 240–241
Alternation, 11, 13, 15, 18, 20, 72; in English, 12, 25, 28–29; laws of, 13, 16, 34, 70, 81–82; of terms, 114, 129–130
Alternational normal form, 72–73, 76
Ambiguity, 1, 4, 27, 55–57
Ancestor, 292–293
And, 10, 28–29, 53–54
Antecedent, 21, 54
Antinomy, 289, 301, 303
Any vs. every, 97, 144
Apagoge, 194
Application of logic, 4–5, 45, 53, 195, 303; to circuits, 77–78; to axioms, 85, 87, 88, 271–273
Aristotle, 108
Arithmetic, *see* Number
Associativity, 10, 13, 67
Attribute, 94, 262, 289
Aut, 12
Axiom, 85–89, 228, 234, 235; of identity, 271–273; of set theory, 302
Behmann, H., 127, 132, 154
Bernays, P., 181, 189, 218, 288; on Löwenheim's theorem, 210–212, 216
Biconditional, 25–26, 35, 55, 61, 68–69; and equivalence, 60, 67; pure, 66; of terms, 128, 130
Boole, G., 20, 73, 132
Boolean algebra, 114, 130–132
Boolean schema, 114–116, 137, 153
Bound variable, 133, 134, 145, 146, 181
Brouwer, L. E. J., 203

Cantor, G., 131, 210, 212, 299
Capture, 161–162
Carnap, R., 147, 217n
Carroll, L., 110
Categorical, 95, 138, 195
Causal connection, 23, 24, 55
Cayley, A., 299
Cd, 242–244, 246
Cellular, 125–127
Church, A., 214–216, 319
Circled numeral, 181
Circuit, electric, 77–78
Circular reasoning, 200
Class, 132, 261; as theme of mathematics, 5, 300; as extension, 94; empty, 94, 130; deferred, 94, 131, 135; algebra of, 114, 130–132; as value of variable, 131, 291; virtual, 136, 291–292; identity of, 289, 290–291; reified, 289–290, 301–303; utility of, 292–295, 300. *See also* Abstraction
Clausal generalization, 227–231
Clause, 72, 83; relative, 133–136, 175, 290
Closed, 134, 147, 153

Closure, 154
Coextensiveness, 128–129, 283, 289–291
Collision, 150, 181
Colon, 50n
Combinator, 286–288
Commutativity, 10, 13, 55, 66. *See also* Symmetry
Compactness, 203
Completeness, 214; in truth functions, 18–20, 86, 87; in subtraction, 88; in quantification theory, 190, 207–208, 215–216; in number theory, 216–218; in identity theory, 273
Computer, 78
Concretion, 134
Conditional, 21–26, 34–35, 54–55; generalized, 22, 95, 139; contrafactual, 23; and implication, 46–47, 49–51; under negation, 68–69; of terms, 114, 129–130; existential, 121, 125; EI, 192; conditional proof, 242
Conditionalization, 242–244, 246
Confinement of quantification, 142–143, 147, 151–154
Conjunction, 9–11, 17–18, 28–29, 34, 53–54; in English, 10, 28–29, 53–54; of terms, 114, 129–130; infinite, 203–205; polyadic, 284, 285, 287
Conjunctional normal form, 83
Consensus, 77, 78
Consequent, 21
Consistency, 39–41, 43, 47; tests of, 41, 73; patent, 43–44; of Boolean schemata, 114–119; in quantification theory, 154, 190, 215–216
Constructivity, 205, 210
Contrafactual, 23
Conversion, simple, 99. *See also* Inversion
Cooley, J. C., 11n, 244
Copula, 129, 135–136, 290–291
Correlation, 298
Curry, H. B., 288

Davis, M., 218
Decision procedure, 88, 112, 172–173, 214, 217; for Boolean schemata, 122–127; for Boolean algebra, 130; for monadic schemata, 154–159, 187, 235; for pure existentials, 184–186, 214; for pure universals, 213–214; impossibility of, 214–216; for finite validity, 215; for real numbers, 218
Dedekind, R., 5, 294, 303
Deduction, 238–239, 245–246; natural, 244; theorem of, 244; soundness of, 246–247; strategies of, 248–254
Degree, 284
DeMorgan, A.; 73, 120, 132, 173, 299; laws of, 14n, 68, 73, 81–82, 84, 140
Denotation, 94. *See also* Extension; Naming; Reference
Description, 274–280; elimination of, 281–282
Designation, 94. *See also* Naming
Development, 73, 83
Dilemma, 227, 254
Disjunction, 15, 74
Distributivity, 70, 82; of ∃, 119; of ∀, 129; of quantifier, 139, 140, 150; of negation, *See* DeMorgan, laws of; Passage
Dot, 29–30
Dreben, B., 208, 222, 224, 233
Duality, 79–84, 224–225
Duns Scotus, 20
Dyadic, 137, 167–168, 187–188

E, *see* Categorical
∃, 116, 119, 285
EG, *see* Generalization
EI, *see* Instantiation
Einstein, A., 195
Empty class, 94, 130, 294
Empty universe, 116–117, 120, 126
English, *see* Paraphrasing
Enthymene, 199–202
Epistemology, 1–4, 276, 283
Epsilon, 135–136, 290, 300
Equivalence, 61–62, 67, 74; and biconditional, 60, 67; under duality, 82, 84; of Boolean schmata, 118, 122. *See also* Interchange

Equivocation, 4, 55–57
Every vs. any, 97, 144
Everything, 138–139, 171–172
Exclusive, 11–13, 18
Existence, 116, 262–264, 282. *See also* Class
Existential, pure, 183–184
Existential conditional, 121, 125
Extension, 94, 145
Extensionality, 289

Fell swoop, 48–49, 63, 82
Finished, 242
Finite universe, 140, 171, 215–216
Flag, 241–242, 244
Foundational system, 87
Free variable, 134, 145, 146, 181
Frege, G., 120, 236; on number, 5, 294, 299; on truth functions, 20, 26, 39, 89; on quantification, 141, 173, 194
Fridshal, R., 78
Full swap, 82
Full sweep, 49
Function, 219–220, 288, 300
Functional normal form, 219–224, 232–234
Functor, 129, 283–287

Geach, P., 293, 296n
Generalization (in proofs), 227–231, 234, 235, 238–242
Generalized conditional, 22, 95
General term, 93–94, 131–133
Generic, 192, 227
Gentzen, G., 244
Gergonne, J. D., 26
Ghazala, M. J., 78
Gödel, K., 136, 208, 216–218, 273
Goodman, N., 23n
Grouping: of symbols, 10, 13–14, 29–32, 143; of component sentences, 14, 27–29, 57–58; of quantification, 139, 142–144, 169, 196

Halmos, P., 288
Herbrand, J., 244; on monadic logic, 127, 147; polyadic, 189, 194, 228, 232
Hilbert, D., 15, 120, 181, 218; on Löwenheim's theorem, 210–212, 216
Homogenizing, 287

Idea, 263–264, 267
Idempotence, 11, 13
Identity, 130, 268–273; of classes, 289, 290–291
If, *see* Biconditional; Conditional
Image, 298
Implicant, prime, 76–78
Implication, 4, 46, 88, 176; and conditional, 46–47, 49–51; laws of, 47, 62–63; tests of, 48–49; truth-functional, 49, 62–63, 82; Boolean, 114, 118, 119; polyadic, 172–173; proof procedure for, *see* Deduction
Inclusion, 128–129, 132, 283, 290–291
Incompleteness, *see* Completeness
Inconsistency, *see* Consistency
Independence, 86–87
Indirect proof, 194
Induction, mathematical, 247, 297
Inference, 45–46, 93; rules of, 85, 87, 228–230, 234, 235, 239; singular, 107, 259, 262; truth-functional, 238
Infinity, 140; lemma, 203; schema, 215–216
Instance, 180, 190, 192
Instantial variable, 190–192
Instantiation (in proofs), 190, 205–206, 213, 220, 238–240; restraints on, 191, 221, 226, 230–231, 240–242; multiple, 220–221, 233–234. *See also* Instance; Substitution
Interchange, 63–65, 122, 156
Interpretation: of sentence letters, 33–34; of term letters, 114–117, 173, 209–211; bypassed, 212
Intersection, 130
Inversion, 285–287, 298, 300. *See also* Commutativity

Jaśkowski, S., 244
Jevons, W. S., 131
Jungius, J., 168

Kaplan, D., 293
Kleene, S. C., 214n, 217, 218
König, J., 203
Korselt, A., 173
Kuratowski, C., 299

Language, *see* Paraphrasing
Leibniz, G. W. von, 132
Lemma, 111
Leśniewski, S., 66
Letter, *see* Schematic; Sentence
Lewis, C. I., 109
Lexicon, 220–221
Limit, 170
Literal, 172
Logic, 1–3, 5, 53, 88
Logical truth, 4–5, 42
Löwenheim, L., 127, 209–212, 216
Łukasiewicz, J., 20n, 26, 30–32, 39, 73; axioms of, 85, 87–89, 236

Main method, 190–193, 205–206
Martin, R. M., 136
Mathematics, 2, 3, 5, 300, 302–303
Matrix, 183. *See also* Open
Meaning, 1, 4, 263–267, 276–277
Mechanical test, *see* Decision procedure
Membership, 135–136, 290, 301
Metamathematics, 218
Mills, B. E., 77, 78
Minkowski, H., 195
Mixed schema, 142, 153, 156–159
Modus ponens, 85, 87, 234, 235
Monadic, 93n, 137, 147, 153, 187; validity, implication, consistency, 176, 235
Müller, R. K., 78

Name, 268, 276–277
Naming, 94, 262–267
Necessity, 2
Negation, 9–10, 15, 34; of terms, 114, 128; of quantification, 137, 144; polyadic, 284, 285. *See also* DeMorgan, laws of
Neumann, J. von, 299, 302–303
New math, 131
Nicod, J., 87, 89
Normal forms: alternational, 72–73, 76; conjunctional, 83; functional, 219, 224
Not, *see* Negation
Notion, 136
Number, 87–88, 209–210; real, 210, 218, 300, 303; elementary theory of, 211; defined, 295–299

O, *see* Categorical; Zero
Object, *see* Abstract object; Existence; Ontology; Physical object; Possible object
Ockham, W. of, 73, 84
Only if, 51, 55
Ontology, 263–267, 276, 282
Open, 134, 145, 154
Or, *see* Alternation
Ordered pair, 167, 297, 299
Ordinary language, *see* Paraphrasing
Ostension, 276

Padding, 285
Pair, 167, 297, 299
Paradox, 289, 301, 303
Paraphrasing, 198; into truth functions, 53–55; inward, 57–59, 196–197; of implicit term, 95–97, 133, 161; of singular term, 107; of Boolean equation, 130; into quantification, 138, 139, 195
Parentheses, 9, 10, 13, 29–30, 143
Particular, 95
Passage, 142–144, 147, 148
Patent, 42–44, 61n
Peano, G., 15, 32, 120, 135–136, 282
Peirce, C. S., 15, 26, 39, 173, 294
Permutation, *see* Inversion
Petitio principii, 200
Petrus Hispanus, 20
Philo of Megara, 26
Physical object, 1, 259, 269
Plato, 167
Polyadic, 167–168, 175–176, 283–285, 297–298
Possible object, 266–267
Post, E. L., 39, 89
Postfoundational system, 87
Predicate, 105, 181

Predicate functor, *see* Term functor
Predication, 132–135
Prediction, 2
Premise, 53, 190–191; tacit, 199–202; temporary, 242–244; descriptional, 275
Prenex, 148–149, 154, 222, 243; strategies of, 150, 184–185, 214
Presburger, M., 217
Prime implicant, 76–78
Product: arithmetical, 296; Cartesian, 298. *See also* Conjunction; Intersection
Pronoun, 133, 146
Proof procedure, 214; for truth-functional validity, 86; for quantificational validity, 173, 190, 224–228, 232–235; for implication, 190; for inconsistency, 190–193, 205–206; for consistency, 215–216. *See also* Completeness; Deduction
Proof theory, 218
Property, 94, 289
Propositional function, 147
Pure: biconditional, 66; quantification, 151–154; existential, 183–185; universal, 213–214

Quantification, 137; under negation, 137, 144; distribution of, 139, 140, 150; scope of, 139, 142–144; in finite universe, 140, 171; history of, 141; confinement of, 142–143, 147, 151–154; purification of, 151–154; stacked, 153, 169, 187; numerically definite, 284, 289
Quantificational: schema, 147, 172; validity, implication, consistency, 172, 173
Quimby, D. P., 203
Quotation, 50, 146, 268

Reductio ad absurdum, 190, 194, 254
Redundancy, 75–76, 84
Reference, 282, 289
Reflection, 285
Reflexivity, 187–189
Relation, 167, 176, 297–298
Relative clause, 133–136, 175
Relative term, 137, 167
Relettering: indifference of, 145, 146; to avoid collision, 148, 150, 193–194, 230, 240–241; for distribution, 150, 312
Resolution, 34–35, 37–38, 61n, 115
Resultant, 298
Rhinelander, P. H., 21n
Right angle, 30n
Rigid routine, 205–206, 213
Rogers, H., Jr., 218
Rosser, B., 244
Russell, B., 136, 141, 236, 276, 282–283; and Whitehead, 15, 32, 51, 89, 141, 147, 303; paradox of, 301, 303

Samson, E. W., 77, 78
Schema, 33, 112, 153, 160; normal, 72–73, 76, 83, 219, 224; term, 114; Boolean, 114–116; existence, 116; mixed, 142, 153, 156–159; quantificational, 147, 172; monadic, 147, 187. *See also* Substitution
Schematic letter: for sentence, 33, 145; for term, 93, 112, 137, 300–301; polyadic, 168, 172, 284–285; vs. variable, 289
Schönfinkel, M., 189, 288
Schröder, E., 20, 39, 74, 84
Science, 1–3, 303
Scope of quantifier, 139, 142–144, 169, 196
Sentence, 131, 160, 176
Sentence letter, 9, 33, 145, 153, 156, 162
Set, 131, 302. *See also* Class
Shannon, C. E., 77
Sheffer, H. M., 20, 87
Shoenfield, J., 218
Simplification, 3, 65–66; of normal schemata, 72, 74–78
Singular term: in inference, 107, 259, 262; and variable, 135, 260–262; vs. general, 259, 261, 288–289, 303; lacking reference, 262–266, 278–280; as description, 276–277; elimination of, 281–282

Skolem, T., 154, 194, 208–211, 221
Smullyan, R., 218
Something, 138–139, 171–172
Soundness: of main method, 192–193; of Herbrand's method, 226, 231; of deductive method, 246–247
Stacked quantifiers, 153, 169, 187
Star, 241–243
Statement, 1–2, 9, 42, 49, 94, 134; singular, 107
Stencil, 181
Stoics, 20
Strategy: in truth functions, 38, 40, 307; in paraphrasing, 57–59, 196–197, 198; in prenexing, 150, 184–185, 214; in deduction, 248–254
Subject, 105
Subscript, 220–221, 224
Substantive, 93, 167; spurious, 138–139, 171–172
Substitution: for '*p*', '*q*', etc., 44, 47, 62, 85, 87; for '*Fx*', etc., 160, 162; for '*Fxy*', etc., 175–181
Substitutivity, 63–65, 122, 156, 271–273
Subtraction, 87–88
Such that, 133–136
Sum, 296. *See also* Alternation; Union
Superstitution, 235
Syllogism, 102–108
Symmetry, 187–188

Tarski, A., 12, 32n, 89n, 218, 244, 288
Telescoping, 27–28
Tense, 96, 195–196
Term: general, 93–94, 131–133; major, middle, minor, 105; absolute vs. relative, 137, 167; polyadic, 167. *See also* Abstraction; Singular term
Term functor, 283–287
Term schema, 114, 177
TF, 238
Theophrastus, 108
Thomas, J. A., 16n
Time, 57, 96, 195–196, 269
Trachtenbrot, B. A., 216, 218
Transitivity, 187–189; of implication, 47; of equivalence, 62
Translation, *see* Paraphrasing
True of, 94
Truth, 1, 9; logical, 4–5, 42; truth function, 16–20, 33, 114–115, 128; truth value, 16, 20; truth-value gap, 21, 79; truth-functional schema, 33; truth-value analysis, 37–38, 41, 43, 61, 65, 115, 157; truth table, 38–39; truth-functional validity, implication, equivalence, 49, 62–63, 176; truth-functional inference, 239
Turing, A. M., 214–216

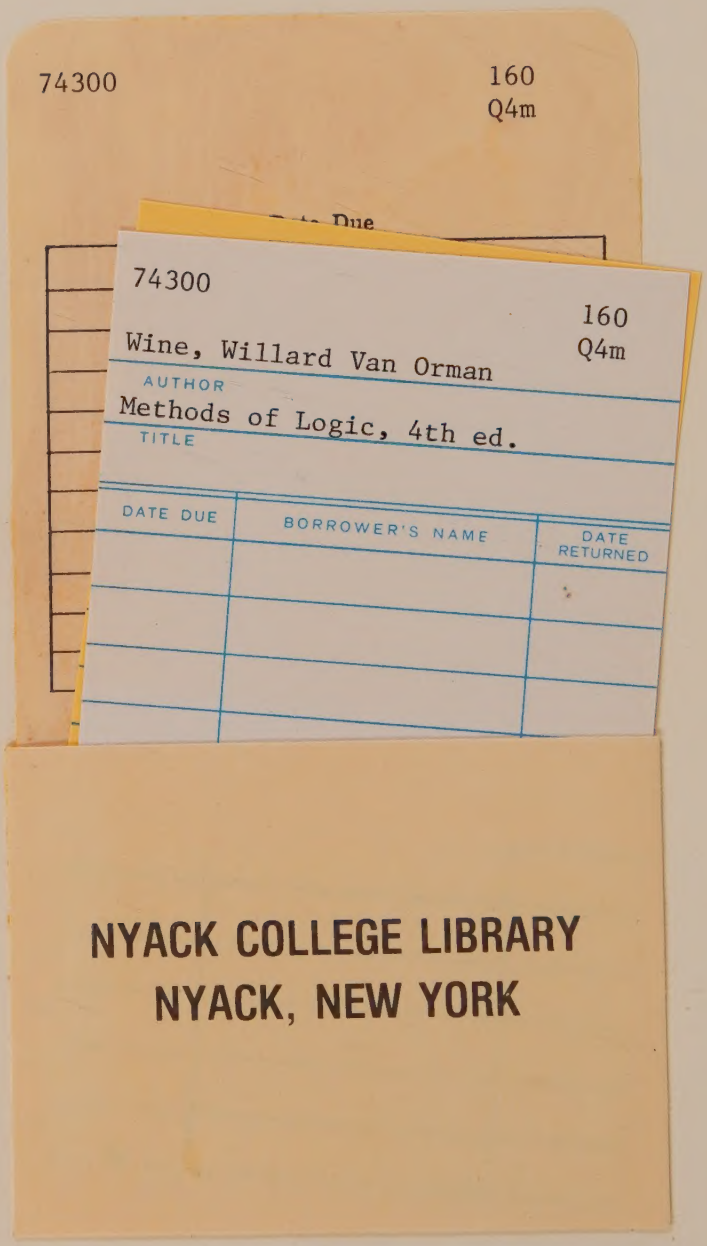
UG, *see* Generalization
UI, *see* Instantiation
Ullian, J. S., 113
Union, 130
Universal: affirmative and negative, 95; closure, 154; pure, 213–214
Universe: of discourse, 111, 115, 120, 198; empty, 116–117, 120, 126; universe class, 130; finite, 140, 171, 215–216
Unless, 55
Usage, *see* Paraphrasing
Use vs. mention, 50, 146n, 268

V, Λ, 130, 141, 294
Validity, 4; truth-functional, 39–42, 49, 176; patent, 42–44, 61n; transmission of, 44, 47, 62, 64, 177, 229–230; under duality, 82; of syllogism, 102–103, 106–107; Boolean, 114–118, 121–122; of open schema, 154; monadic, 154–159, 176, 235; quantificational, 172–173, 190, 212; of pure existential, 183–184, 189; proof procedures for, 190, 224–228, 232–235; in identity theory, 271
Van Heijenoort, J., 173n
Variable, 133, 134, 137, 146, 181; and singular term, 135, 260–262; instantial, 190–192; elimination of, 284–288; vs. schematic letter, 289. *See also* Relettering
Vel, 12
Venn, J., 98, 103, 109–110

Vinculum, 32
Virtual, 136
Von Neumann, *see* Neumann

Weierstrass, K., 5
Whitehead, A. N., 270n; and Russell, 15, 32, 51, 89, 141, 147, 303
Wiener, N., 299, 303
Wittgenstein, L., 39
Words, *see* Paraphrasing

Zermelo, E., 301–302
Zero, 296, 299


74300
160
Q4m
74300
160
Q4m
Wine, Willard Van Orman
AUTHOR
Methods of Logic, 4th ed.
TITLE
DATE DUE
BORROWER'S NAME
DATE RETURNED
NYACK COLLEGE LIBRARY
NYACK, NEW YORK

Nyack College
160 Q4m
Quine, W. V. (W - Methods of logic / W.V. Quine.
3 6847 00032 2240